Exhibit G

Who Facilitated Misreporting in Securitized Loans?

JOHN M. GRIFFIN and GONZALO MATURANA*

ABSTRACT

This paper examines apparent misrepresentation among securitized non-agency loans using three indicators: unreported second liens, owner occupancy misreporting, and appraisal overstatements. We find that approximately 30% of loans exhibited some indication of misrepresentation. Misreporting is similar in both low and full-documentation loans and is associated with a 51% higher likelihood of delinquency. Interest rates at origination and activity around securitization thresholds indicate that originators were largely aware of second-lien misreporting and, to some extent, inflated appraisals. Misrepresentation also explains substantial cross-sectional differences in future MBS losses. Losses were predictable and initiating from the practices of MBS underwriters and loan originators.

---

*Griffin and Maturana are at the University of Texas at Austin. We are grateful to Kenneth Singleton (the editor), an anonymous associate editor, two anonymous referees, Andres Almazan, Aydogan Alti, Malcolm Baker, Bo Becker, Jonathan Cohn, Shawn Cole, Josh Coval, Giovanni Favara, Samuel Hansen, Victoria Ivashina, Alexander Ljungqvist, Chris Malloy, Jordan Nickerson, Amiyatosh Purnanandam, Jonathan Reuter, Alessio Saretto, Amit Seru, Clemens Sialm, Erik Stafford, Sheridan Titman, as well as seminar participants at Arizona State University, Boston College, Harvard Business School, the University of Chile, the University of Massachusetts Amherst, the University of Texas at Austin, and the 2013 CAREFIN Bocconi Banking Conference for helpful comments. We also thank Kevin Mei, Jacob Rachiele, and especially Christopher Miller for excellent research assistance. Griffin is now a consultant and involved in legal actions related to mortgage fraud. The analyses and conclusions of the paper are unaffected.

At the heart of the recent financial crisis was an extremely rapid deterioration in the value of non-agency Mortgage-Backed Securities (MBS) and Collateralized Debt Obligations (CDO) derived from MBS. To fully understand the crisis, one needs to thoroughly understand the nature and incentives surrounding the ground-level collateral that later filtrated into many types of structured products. Explanations for the extremely poor loan performance include a decrease in loan quality (Demyanyk and Van Hemert (2011)), an exogenous housing bubble (Foote, Gerardi, and Willen (2012)) or decrease in housing prices (Mayer, Pence, and Sherlund (2009)), and the poor incentives tied to the "originate-to-distribute" model (Keys, Mukherjee, Seru, and Vig (2010), Purnanandam (2011), and Keys, Seru, and Vig (2012)). Less work has focused on the role of mortgage misrepresentation and its relationship with underlying loan performance.

Mortgage fraud is a source of interest for the financial press and an area of increased focus for the FBI (Federal Bureau of Investigation (2011)). Yet, it is difficult to detect from prosecuted cases and the press if these are rare events or widespread patterns, and who are the responsible parties. This paper examines the extent and role that borrowers, appraisers, originators, and MBS underwriters may have played in three different types of mortgage misrepresentations and their effects on MBS losses.

A matching algorithm allows us to link large datasets of non-agency MBS loan data from 2002 to 2007 with county-level official transaction information and perform detailed loan monitoring along three dimensions. First, we examine the prevalence of first-lien loans that are recorded in MBS loan-level data as having only a first lien, yet when matched to county-level information show a second-lien loan issued simultaneously to the first-lien loan. This form of misrepresentation is the least debatable evidence, since we are comparing data from servicer/trustee reports to what is recorded in tax records. Second, we examine cases where loan-level MBS data indicates that a house is owner occupied, and yet county-level data shows that the tax records are sent to a different, non-business address. Third, we examine appraisal values that appear to be overstated. We use an industry-leading

automated valuation model (AVM) that provides a statistical valuation for a house at the time of loan origination. Given that the approach relies on a valuation model that will have estimation error, it is an empirical question as to whether the AVM or the appraisal is more accurate. To be over-conservative, we flag any transactions occurring at appraisals more than 20% above the AVM as potentially inflated, and we further investigate the causes behind this measure.

We find that 13.4% of loans reported as having no second lien (10.2% of all loans) do have a second lien. Approximately 7.7% of loans marked as owner occupied (6.7% of all loans) may not be owner occupied. Additionally, even after using our over-conservative measure of inflated appraisals, 17.8% of homes have appraisals that appear inflated. Since we show that the inflated appraisal metric is far too conservative and we only examine three metrics of misreporting, our overall frequency of misreporting (30.1%) might best be considered as a lower bound for the amount of non-agency mortgage misrepresentation. Aggregate misreporting frequencies are similar for low and full-documentation loans (29.6% and 30.6%, respectively), suggesting that misreporting is not simply due to lack of information disclosure. Loans with an unreported second lien, owner occupancy, or appraisal overstatement indicator are 97%, 8%, and 34% more likely to become delinquent than loans with no misreporting indicators, respectively. The findings are also strong for a measure of direct default and robust to the inclusion of Core Based Statistical Area-origination quarter fixed effects and many other loan-level controls.

We then examine the role of borrowers, appraisers, originators, and MBS underwriters in misreporting. Loans with unreported second liens are originated with higher interest rates than loans correctly reported. These rates are similar to (though slightly higher than) the interest rates for first-lien loans with reported second liens, indicating that originators were seemingly aware of and accounted for the second-lien risk. Additionally, in more than two-thirds of the cases of unreported second liens, the same originator issued the first and second lien; originators should have been aware of the second lien. We find that the fraction of

unreported second liens jumps around securitization thresholds, suggesting that originators misreported second liens with intentions of ultimate loan securitization. We examine the possibility that loan originators accurately reported second-lien data to underwriters, who later omitted this information at issuance. However, in explaining unreported second liens, both originator and underwriter fixed effects are economically important.

Owner occupancy misreporting is not accompanied by a materially higher interest rate, indicating that originators may have been largely unaware of buyers' intended usage of the property. Loans with inflated appraisal values sold with some premium, indicating that loan originators viewed these loans as riskier. Owner occupancy and appraisal overstatement misreporting for the most part do not jump around securitization thresholds or vary much across originators, indicating that the practices were not driven by securitization.

In contrast, second-lien misreporting varies considerably across originators, whether standalone or part of a large corporation. Originators with high levels of second-lien misreporting also have abnormally poor loan performance even after controlling for the individual loan-level misrepresentation indicators. This suggests that our misrepresentation indicators may not be capturing the full extent of misrepresented loans or some other aspect of poorly performing originating practices that are correlated with mortgage misrepresentation. The prevalence of misreporting by originator and originator loan performance is fairly persistent over time. In contrast, second-lien misreporting increases from 2002 to 2005 with the growth of the non-agency market. This rise is primarily related to a gain in market share by originators that exhibited similarly high levels of second-lien misreporting even back in 2002. Second-lien misreporting peaks around 14% in early 2005 and drops to 0.7% in the third quarter of 2007 as originators with the highest levels of second-lien misreporting went out of business. Additionally, originators with both good and poor loan performance seem to be affected by the incentive to securitize around credit score thresholds, although the jumps around securitization thresholds are much larger for poor performing lenders.

The strong relation between the appraisal indicator and future delinquency indicates that

the AVM statistical model was substantially more accurate than most appraisers but does not speak to whether the poor appraiser performance was due to random appraiser error or appraisers catering to loan officers. To examine these differences, we focus on refinances, where the value of the transaction is purely set by the appraisal. Appraisal overstatement is 74% more likely to occur on refinances than on purchases. 49% of appraisals used for refinance cluster exactly on loan-to-value ratios in five-unit increments, and these have consistently higher levels of appraisal overstatement as well as future delinquency rates. This cluster around bank threshold loan-to-value ratios suggests that appraisers target numbers from loan officers.

While we show that loan originators were likely aware of second lien misrepresentation, this raises the question of what bank underwriters knew. Monitoring services such as Clayton Holdings provided loan monitoring for a sample of loans for each individual MBS. These loan monitoring services seem to include the three types of misreporting that we study, among others. This suggests that MBS underwriting banks knew that some of the MBS representations at issuance were apparently incorrect. When examining misreporting at the pool level, we find that the percentage of potentially misreported loans in MBS pools is one of the most important cross-sectional determinants of subsequent MBS performance. Occupancy misreporting and appraisal overstatements are somewhat constant across underwriters, but second-lien misreporting is not. The amount of second-lien misreporting varies considerably across underwriters, even after controlling for their choice of loan originators. Extremely poor MBS performance was not just a function of disclosure or bad luck; some information available even before MBS issuance predicted their future demise.

Our paper adds to a growing literature examining mortgage misrepresentation. Jiang, Nelson, and Vytlacil (2013) find evidence of income falsification in low-documentation loans at a large bank. Garmaise (2013) finds that borrowers at a large bank who report unverified assets slightly above round number thresholds are more than twice as likely to default.[1]

[1]Carrillo (2011) examines how to detect mortgage misrepresentation, including inflated appraisals, in a large county. Amromin, Huang, Sialm, and Zhong (2013) find that complex mortgages attract sophisticated

Ben-David (2011) finds evidence of inflated appraisals and 'cash back' deals in highly-levered deals in the Chicago area. Haughwout, Lee, Tracy, and Van der Klaauw (2011) focus on the role of real estate investors in the housing bubble and show that by misreporting occupancy status, investors were able to increase leverage and contribute to higher rates of default. In a parallel study, Piskorski, Seru, and Witkin (2013) use credit agency data with loan-level data from 2005 to 2007 and detect that around 10% of non-agency loans exhibit owner occupancy and second-lien misreporting; this is later associated with 60% higher probability of default, yet misreporting was not reflected in MBS pricing or subordination. The fact that our aggregate levels of misreporting are higher is explained by our inclusion of appraisal overstatements.[2] It is comforting that although the papers use entirely different data sources, they reach similar conclusions concerning the existence of large-scale misreporting and its effects on loan performance. Nevertheless, the papers have different focuses regarding misreporting. Piskorski, Seru, and Witkin (2013) find some regional patterns in misreporting, but little relation between misreporting and reputation, top management compensation, or regulatory environment. We use one additional measure of misreporting (over a longer period) and focus on the various roles of borrowers, appraisers, originators, and underwriters in misreporting. We show that misreporting indicators are economically important cross-sectional determinants of future MBS performance, but we do not examine what was known by investors. Piskorski, Seru, and Witkin (2013) show that the effects of misreporting were not reflected in MBS pricing or subordination; our combined inferences suggest that MBS investors were unaware of, yet adversely affected by, misreporting.

In addition to understanding the non-agency loan market, our paper is related to a growing literature on improving the understanding of problems in structured finance in the period prior to the recent crisis. Ashcraft, Goldsmith-Pinkham, and Vickery (2010) show

borrowers who default strategically. Jiang, Nelson, and Vytlacil (2014) show how lax screening can also hurt the originating bank. Agarwal and Ben-David (2013) study the role of loan officers' incentives in the crisis and conclude that they adversely affected underwriting standards.

[2]We find practically the same levels of second-lien misreporting when including home equity lines of credit (HELOCs) (13.4% compared to 13.6%) and a slightly higher level of occupancy misreporting (7.7% compared to 6.4%).

that RMBS subordination standards deteriorated from 2005 to 2007. For CDOs, Griffin and Tang (2011, 2012), and Griffin, Nickerson, and Tang (2013) show rating agencies issuing inflated ratings, upward adjustments in credit rating models increasing from 2004 to 2007, and rating agencies seeking to match their competitors ratings to please issuers. Begley and Purnanandam (2013), Demiroglu and James (2012), and Faltin-Traeger and Mayer (2012) find that MBS structuring may reflect considerable private information within the market. The above literature, like our findings, indicates considerable private information across deals and largely cuts against the view that the crisis was a completely random event unrelated to structuring incentives.

The rest of the paper is organized as follows: Section I provides an overview of the non-agency MBS market and relates some of its features to the theoretical literature on misreporting. Section II describes the data sources, with special emphasis on matching. Section III introduces the three potential misrepresentation and loan performance measures. Section IV studies the effect of misreporting on loan performance. Section V investigates the role of borrowers, appraisers, loan originators, and MBS underwriters in the facilitation of potential misreporting. Section VI further looks at misreporting across originators. Section VII studies the impact of loan misreporting on future MBS losses, and Section VIII briefly overviews our findings and draws implications for the future.

I. Mortgage Origination Background and Theory

In this section, we first survey some specific aspects of the non-agency MBS market and then relate to the theoretical literature on misreporting.

A. Securitization Market Background

The non-agency securitization market grew rapidly from roughly $85 billion a year in 2000 to $1.1 trillion a year in 2005 and 2006 (Goodman et al. (2008)).

Fundamental features of the securitization market are asymmetric information and the

potential for agency conflicts (Riddiough (1997), Bernardo and Cornell (1997), and DeMarzo (2005)). Consequently, the securitization market also has a number of checks to assure asset quality among the many different players. A brief outline of potential conflicts is as follows. Homeowners might want to misreport features of their application to obtain better loan conditions. However, the loan officer collects and verifies key pieces of information (such as income, credit history, and other debt obligations). On the other hand, the mortgage broker himself is often compensated on loan volume. Yet, originating banks can fire or ban loan officers with large numbers of inaccurate loan files. Banks also depend on independent appraisers to verify home valuations. To maintain their licenses, appraisers are required to issue unbiased and accurate appraisals. Yet, since appraisers often depend on loan officers for current and future business, they also have a conflicting incentive to please the loan officer. Underwriters realize they purchase loans from originators that have more information about the loans. To receive higher prices for loans, originators might have an incentive to inaccurately report the information they provide to underwriters. However, to monitor this conflict, underwriters hire loan monitoring companies who perform detailed examination of securitized loan pools. Structured Finance Underwriters make larger profits if the pools are aggressively structured. Yet, underwriters have an incentive to uphold their reputation and correctly represent deals. The securitization chain consists of considerable asymmetric information but also an established system of checks and balances along the way. Each member of the securitization chain largely depends upon the next level for its future business. Most channels flow directly or indirectly up to the incentives of the underwriting banks.

B. Theory Overview

Becker (1968) and Ehrlich (1973) model the rationality of crime and how illegal activity occurs if potential gains exceed enforcement penalties. Additionally, in the typical "Inspection Game," misreporting happens with a probability increasing in monitoring costs. These views contrast with the theoretical extremes where fraud either never occurs (Grossman and

Hart (1980)), or it never catches investors by surprise (Akerlof (1970)).

Bolton, Freixas, and Shapiro (2007) show that when profit margins are high, banks will have a stronger incentive to misreport. Povel, Singh, and Winton (2007) model how fraud is more prevalent and less detected during boom times due to lax monitoring. Misreporting the leverage in a house would have a minimal effect on losses as long as house prices are increasing rapidly, but the extent would be revealed in a burst. The conventional view of reputation is that the loss of future business is enough to induce market participants to correctly report (Booth and Smith (1986) and Chemmanur and Fulghieri (1994)). Griffin, Lowery, and Saretto (2014) show how this conventional wisdom can hold with simple transparent securities but breaks down with complex structured products. With structured products like MBS, ABS, and CDO, investors will only learn the true value of the underlying assets in the next market downturn. This ability to misreport without being detected for an extended period with structured products, unlike simple securities, can create the incentive for even a reputable underwriter to disregard their reputation and misrepresent securities.[3]

If investors are performing their own loan-level analysis of potential fraud with the same level of information as loan officers, then there is little incentive to misreport. Yet, to assume investors would have the access to similar loan-level information available to loan officers is inconsistent with the key nature of asymmetric information in the MBS market (Riddiough (1997) and DeMarzo (2005)). Additionally, pension funds, insurance companies, and some banks are said to have partially relied upon a security's credit rating. Second liens, owner occupancy, and loan-to-value were key inputs in credit rating models, and rating agencies seemed to have accepted loan-level data as accurate.[4] If investors were purchasing based on rating agency and investment bank certification, then there is a potential incentive to misreport. Since the theoretical literature above outlines the possibility of costly

[3]Similarly, Mathis, McAndrews, and Rochet (2009) and Fulghieri, Strobl, and Xia (2014) argue that rating agencies may strategically burn their reputation and Bolton, Freixas, and Shapiro (2012) and Sangiorgi and Spatt (2013) show that the rating agency incentives to cater to banks increases with security complexity.

[4]The president of Clayton claims that credit agencies were not interested in using their monitoring services because this new information would yield lower ratings and cost them future business (Financial Crisis Inquiry Commission (2011), pages 167-168).

misreporting, it thus becomes an empirical question as to whether this behavior occurred, and if so, what role might be attributed to the various players in the securitization chain.

II. Data and Measures

The data for this study comes from four main datasets: Lewtan's ABSNet and HomeVal datasets, along with DataQuick's Assessor and History Files. The first two Lewtan datasets provide loan-level and home valuation information, while DataQuick provides house characteristic and transaction information. Lewtan is one of the industry leaders in providing performance metrics and origination information on the mortgage loans that back US non-agency MBS. Lewtan compiles and cleans loan-level information as reported in non-agency MBS raw servicer/trustee loan-level data tapes, like the ones available in free writing prospectus documents. Lewtan's ABSNet Loan contains information about more than 18 million residential loans that were issued either for the purchase or refinancing of properties between January 2002 and December 2011 and provides origination information such as the appraised value of the property, the detail of the documentation provided by the borrower, the purpose of the property as reported by the borrower (owner occupied, second home, or investment), the loan amount, the loan-to-value ratio, the interest rate, the credit score of the borrower, and the origination date. Additionally, the database provides the history of payments for each loan and other metrics. HomeVal, on the other hand, provides home valuations at the time of origination based on Lewtan's ABSNet proprietary automated valuation model (AVM). The AVM is developed by Collateral Analytics, a firm that specializes in AVM models and mortgage risk tools.[5]

DataQuick is one of the largest providers of real estate data in the US. DataQuick's Assessor File holds detailed information on residential properties as registered from County

[5]The quality of their valuations is supported by AVMetrics, a firm dedicated to evaluating and ensuring the correct compliant use of AVM's best practices. Collateral Analytics' AVM has continuously ranked among the top industry performers in competitions for AVM accuracy. Collateral Analytics AVM value is a weighted combination of several valuation methods including hedonic and regression models, time adjustments, appraisal emulation, and neighborhood and characteristic matching.

Assessors. The database includes more than 120 million properties and provides specific information such as the address and physical characteristics of the property. The History File, in turn, provides information related to residential property transactions. From January 2002 to December 2011, more than 175 million transactions are recorded, involving almost 62 million properties from the Assessor File. The History File gives information on the transfer date of the property, the identities of the buyer and seller, the mailing address of the buyer, and the various loans involved in the transaction, among other information.

A. Merging Process

The richness and the detail of the information in the datasets along with their high coverage are clear advantages when studying mortgage misrepresentation on a massive scale. However, given that Lewtan and DataQuick are two independent and unrelated companies, there is no key variable or unique identifier that links the mortgage loans in ABSNet Loan to the corresponding transaction in DataQuick's History File. The identity of the buyer and the property address are available in DataQuick but not in ABSNet Loan, which only provides the zip code. Nevertheless, firms specializing in loan monitoring or performance often claim they can link these or similar databases with considerable accuracy. Since erroneous matching can lead to potentially overstating misrepresented loans, we take a conservative approach and perform extensive diagnostics on our matching procedure.

We match first and second liens separately. Residential loans are matched with the transactions according to their zip code, loan amount, interest rate type (fixed or adjusted rate mortgage), loan type (conventional, FHA, or veteran), originator, and purpose of the transaction (refinance or purchase). Additionally, the transfer date is required to be within a [-15, 30] day interval around the loan origination date, and the differences in transaction prices, when available, are required to be within $1,000. We only consider a transaction-loan pair a match when it is unique.

The algorithm matches 34.6% of the first liens with an initial loan amount over $30k in

ABSNet Loan. We compare characteristics between matched and unmatched loans (in the Internet Appendix C)[6] and see that both samples are quite similar in loan-to-value ratio, interest rate, credit score, asset type, performance, and most other characteristics. Since transaction-loan pairs with smaller loans are less likely to be unique, matched loans have slightly higher loan values. In terms of geographic distribution, we obtain a higher matching rate in California and Florida, which are also the two more important states in our sample in terms of number of loan originations.

B. Sample Selection

Our transaction data begins in 2002 (prior to which the subprime market was limited) and runs through the end of 2007, after which few non-agency MBS were issued. We consider first-lien loans with an initial loan amount over $30k. Like Ben-David (2011), we also drop loans with an LTV over 103% because values this high are often associated with data errors. To maximize data quality regarding special houses that may be difficult to appraise, and to focus on normal loans, we also drop the loans associated with the largest 1% of the transaction values in each state and loans that were reported as being for homes of more than one unit. To check LTV and combined LTV (CLTV) accuracy, we collected our own sample of MBS prospectuses and compared loan-level data, leading us to drop all MBS deals in which all of the underlying mortgages are recorded with an LTV equal to their combined LTV.[7]

Table I, Panels A through C provide a description of the main sample. After applying the filters above, we end up with 3,143,755 loans.[8] The median loan amount is roughly $234k, the

[6] In the Internet Appendix, Figure IA.2 compares the kernel densities of different variables in both samples, and Table IA.I compares the different matching qualities across loan characteristics, loan performance, and geographic distribution of both samples and the samples used for our analysis, which are constructed as we describe below.

[7] We manually checked 87 free writing prospectuses and found data errors when the LTV equals the combined LTV for all the mortgages in a MBS deal. Consequently, we dropped 1,031 deals where LTV and combined LTV were always equal since they could cause an overstatement of second-lien misreporting. For the rest of the deals in ABSNet where the loans had different LTV and combined LTV values, the free writing prospectuses were similar to the ABSNet data.

[8] We matched 5,284,624 first liens originated during the period from 2002 to 2007 that were used for

median LTV is 80%, and the median credit score is 675. The sample consists of 16.5% prime loans, 47.6% subprime loans, and 36% Alt-A, negative amortization, or scratch-and-dent loans. Additionally, 87% of the loans are reported as owner occupied, with the remaining reported as investments or second homes. The proportion of loans with an adjustable rate is 67.4%, while the remaining have a fixed rate. Low/no-documentation loans are 55.8%, and full-documentation loans are 44.2%. With respect to the performance of the loans in the sample, 33.1% became seriously delinquent (past due 90 days or more) while 26.1% went into foreclosure before July 2012. Also, 8.2% of the loans went into foreclosure and the borrower did not make any payments between the first payment that was missed and the foreclosure date (direct default). Finally, 2.2% of the loans went into early delinquency, meaning that the loan became seriously delinquent within six months from the first payment date.

<<TABLE I HERE>>

We obtain a high transaction-loan matching rate in California and Florida –the two states with higher loan origination rates– possibly due to differences in property transaction information across states. To verify that the overrepresentation of these two states in the main sample does not drive our results, we construct another sample from the main one in which we further drop the loans originated in California or Florida. Finally, because of missing information, it is not always possible to compute all three misrepresentation indicators for all the loans in the main sample. To facilitate comparison and as an extra check on our results, we drop the loans for which we do not have the three indicators from the main sample (we call this the common sample).

III. Misrepresentation Indicators

In this section we explain our misrepresentation indicators.

purchase or refinancing with an initial loan amount over $30k. After applying the additional filters explained above, 5,105,221 loans remain. An additional 1,031,649 loans are lost when we drop the MBS pools where LTV and CLTV values are the same for all MBS in the pool as described above. Finally, we drop loans where any of the variables used as controls in our regressions is missing, leaving us with 3,143,755 loans.

A. Misrepresentation Indicators

Misrepresentation is simply financial misreporting with the intent to deceive. The FBI distinguishes between two types of mortgage misrepresentation: (i) fraud for property/housing, and (ii) fraud for profit. The first type consists of misreporting by the borrower in order to obtain funding to purchase his or her primary property. The second type involves more sophisticated schemes designed to obtain illicit monetary gains from property sales. Methods to facilitate misrepresentation include inflated appraisals, occupancy status misreporting, unreported second liens, property flipping, and the falsification of the borrower's financial information such as bank statements, tax return documents, income, assets, and liabilities.[9] The three indicators of potential misrepresentation that we focus on are unreported second liens (often referred to as 'silent seconds' in the industry), owner occupancy misreporting, and appraisal overstatements. We focus on these indicators because (i) they can be constructed on a large scale from available data, (ii) they are commonly discussed forms of misrepresentation, and (iii) they are listed as indicators used by firms in the loan monitoring industry (such as Clayton). We call the variables 'indicators' because they are constructed to be associated with misrepresentation in general, but may capture loans that appear suspicious but have a legitimate justification. Ultimately, whether the misrepresentation indicator captures misrepresentation is an empirical question. For the first part of our analysis, we focus on misreporting without discussing which of the relevant parties are responsible. Later, we analyze the information that may have been known by buyers, appraisers, mortgage originators, and MBS underwriters. All of our indicators are constructed with information from public records on the closing date of the transaction, meaning our indicators could have, in most cases, been constructed prior to MBS origination.

[9]Other schemes documented by the FBI include foreclosure rescue, loan modification, builder bailout/condo conversion, equity skimming, home equity conversion mortgage, debt elimination, and more.

A.1. Unreported Second Lien

A second lien allows a borrower to take additional debt, giving the borrower less incentive to repay the loans and making the initial debt riskier. Therefore, to evaluate the risk of a first-lien loan, it is important to know whether the borrower has a second lien. We construct the dummy variable *Unreported Second* that takes the value of one if the loan in ABSNet associated with that transaction does not disclose the existence of the second lien (i.e., LTV=combined LTV) but both a first and second lien are recorded in county-level records as captured in DataQuick. To be flagged as potentially misreported, we also conservatively require the LTV of the first lien to be larger than or equal to 80%.

A.2. Occupancy Misreporting

Borrowers who both own and occupy a property are less likely to default than borrowers who do not occupy the property. Consequently, originators charge lower interest rates and require smaller down payments for owner occupants. This in turn gives borrowers and/or originators the incentive to misreport occupancy status.

Similar to Lee, Mayer, and Tracy (2012), we are able to estimate the occupancy status from the county-level transactions using the DataQuick database. We compare the mailing address (where the county sends the tax bill) to the purchased property address. If the mailing address differs from the property address, then we consider the property as being a second home or an investment property. Some people might have their taxes sent to their business address (or a Post-Office Box), so we additionally require that the mailing address neither corresponds to a commercial address nor to a P.O. Box. The variable *Occupancy Misreporting* takes the value of one if the self-reported occupancy status associated to the loan in ABSNet Loan is marked as "owner occupied," but we estimate otherwise from DataQuick's county-level data. Since owner occupancy status is constructed from where the purchaser files to have their first tax bill sent, the measure may capture purchasers who later become owner occupants. Piskorski, Seru, and Witkin (2013) use credit data which notes where a

person is having their bills sent. They define a house as non-owner occupied if bills are never sent to the property in the first twelve months after purchase. Their measure is 86% as large as our measure over their sample period, indicating that a limitation of our owner occupancy measure is that it likely captures a small set of late movers who do not occupy the home immediately after purchase but later become owner occupants.

A.3. Appraisal Overstatement

If the appraiser gives an inflated appraised value for the property, the borrower can secure a larger loan. If the difference between the appraised value of the property and its fair value is large enough, the borrower can obtain a monetary gain at the expense of the lender by defaulting on the mortgage payments (i.e., misrepresentation for profit). Even if the borrower has no intention to default, it lets the borrower put less money down (i.e., misrepresentation for housing).

As a proxy for the fair value of a property at the time of origination, we use Lewtan's proprietary AVM.[10] We define the dummy variable *Appraisal Overstatement* that takes the value of one if the appraised value recorded before origination exceeded the AVM value by more than 20%. In contrast to the other two measures, the AVM originates from models. Both the appraisals and the AVM will have estimation error. Ultimately, it is an empirical question as to whether the appraisals or AVM are more accurate. To the extent that an AVM is more accurate than appraisers, it is indicative that either appraisers made mistakes or appraisals were potentially inflated. We will later try to separate these two possibilities. To be conservative, we use a high threshold of 20% because it seems less likely that an appraiser was more than 20% off from the AVM without an intention to match the number provided by the bank. Empirically, we show in Figure IA.4 that loan performance is considerably worse

[10] It is important to highlight that the AVM value at origination is constructed based on information available at that time, that is, it is not subject to look-ahead bias. Additionally, we consider the AVM model as missing in cases in which the AVM value and the appraised value were exactly the same, as it seems unlikely that the appraised value from a combination of statistical models can exactly coincide with the realized value.

with appraisals 5% larger than the AVM value and that the underperformance increases from 5%-20%, indicating that our 20% filter is too conservative. Nevertheless, we apply it due to the nature of the AVM originating from a model. We define *Misreported* as a dummy variable that takes the value of one if one or more of the three misrepresentation indicators is true.

A.4. Summary Statistics

The mean values of the misrepresentation indicators defined previously are presented in Table II. 10.2% of first-lien loans contain a second lien that is not disclosed. This percentage is 13.4% when considered as a fraction of all loans marked with no second lien. The occupancy misreporting indicator appears in 6.7% of the sample. The most common misrepresentation indicator is appraisal overstatements, which appear in 17.8% of the loans. Aggregating across all indicators, 30.1% of the loans exhibit at least one misreporting indicator. The correlations between the three misreporting indicators are fairly low (3% on average).

Appraisals can be understated as well. Nevertheless, as shown by the red line in Figure IA.4, the distribution of the appraisal overstatements is neither symmetric nor centered at 0%. Inconsistent with random appraisal errors, we only find that 5.4% of the loans show appraisals where the AVM is 20% or more, indicating 12.4% (17.8-5.4) more overstated appraisals than understated.

Additionally, unreported second liens and occupancy misreporting are considerably higher in purchases than in refinances. The opposite is the case for appraisal overstatements. This is interesting since appraisals are the sole determinant of the transaction price with refinances. Lenders would also presumably be aware if the house was owner occupied or had a second lien with a refinance. Interestingly, misreporting does not seem to be a simple function of available information at origination. Both unreported second liens and appraisal overstatements are slightly higher in full-documentation loans than in low/no-documentation loans (1.6% and 0.6% higher, respectively).

<<TABLE II HERE>>

Figure 1 shows the evolution of the different misrepresentation indicators by quarter for the period from 2002 to 2007. The prevalence of unreported second liens increases rapidly and peaks about a year before the top of the housing market in the first quarter of 2005. Owner occupancy misreporting seems to gradually decrease from an average of 8.3% in 2002 to an average of 5.1% at the beginning of 2007. Appraisal overstatement oscillates but remains at high levels throughout the period.

<<FIGURE 1 HERE>>

In the second and third quarter of 2007, the prevalence of unreported second liens plummets. Since reported second lien origination stays at a similar high level, the drop is not due to a drop in second lien origination.

B. Possible Mismatch Rates

We are quite careful in our matching procedure, but mismatches can still occur and could inflate the reported amount of misrepresentation for second-lien misreporting and owner occupancy.

We check the accuracy of our matching and its potential effect on misreporting in three ways. First, as discussed above, second-lien misreporting drops dramatically in 2007 and only reaches 0.7% in the third quarter of 2007. If the mismatch rate were high, one would not expect the proportion of misreporting to be so low in 2007 unless the proportion of second lien volume origination was low. However, the amount of securitized loan originations is similar (with a larger proportion of reported second liens) to that in 2002, when misreporting was 7% on average. Even if we assume that the entire second-lien misreporting in the third quarter of 2007 is due to improper matching, our results still only show that 0.7% of second-lien misreporting is due to improper matching.

Second, we are able to check the matching accuracy using a sample of 6.4% of the matched first-lien loans, where we also match the second lien associated with the same property

through ABSNet. This can be achieved by comparing the loan amount of the first lien with the senior loan amount information associated with the second lien and by comparing FICO scores of both the first and second liens to check if they are within 20 FICO points.[11] Taking a conservative approach, we attribute all mismatches to the first-lien loan rather than to the second-lien loan, and find this potential mismatch percentage to be 5% in our sample. [12]

Third, we are granted access to proprietary loan pools with actual addresses for three MBS deals. Here we check the accuracy of our matches and find that they are 98.2%, 98.5%, and 97.4% accurate. Some of the mismatch could be due to differences in address reporting conventions; nevertheless, assuming these samples are representative suggests a mismatch rate of at most 1.9%,[13] which implies a potential overstatement rate of 0.56% (0.019*0.294) for second-lien misreporting and 0.26% (0.019*0.134) for occupancy misreporting (0.82% combined).

In summary, we use three different methods to examine our matching and find an upper mismatch rate bound for each. If we subtract off the effect on misreporting of the 1.9% potential mismatching found in loan pools with addresses, we obtain an aggregate misreporting of 29.3% (0.301-0.0082).

IV. Does Misreporting Affect Loan Performance?

In the previous section we showed an extremely large amount of potential misrepresentation. However, these measures need external validation. If the indicators are capturing mortgage misreporting effectively, then these loans should underperform loans that are not

[11] Of course, FICOs within 20 points could be different individuals with similar FICOs. There may also be legitimate matches where the FICO score varies by more than 20 points because FICOs from different agencies or points in time are used.

[12] Nevertheless, the effect on our measures is considerably smaller. Only mismatched transactions that have a second lien will contribute to inflate the Unreported Second indicator. Likewise, only mismatched transactions of second-home or investment properties will contribute to inflate the Occupancy Misreporting indicator. From DataQuick, we estimate that 29.4% of our sample has a second lien and that 13.4% of the properties are second-homes or investments. Hence the overstatement rate due to mismatches is 1.47% (0.05*0.294) for second-lien misreporting and 0.67% (0.05*0.134) for occupancy misreporting.

[13] For the three deals with 98.2%, 98.5%, and 97.4% matching accuracy, the number of loans with available addresses are 1,487, 529, and 383, respectively. This gives a loan-weighted matching rate of 98.1%

affected by misreporting or misrepresentation. In most of our tests we focus on serious delinquencies, followed by direct defaults as our measures of loan performance.

A. Summary Analysis

Figure 2 shows loan performance throughout the entire credit score spectrum, separated by low/no-documentation loans and full-documentation loans. Panel A shows that the delinquency effect is strong for second-lien misreporting. Misreporting is strongly related to delinquencies across all ranges of the credit score distribution, but particularly for low credit scores. The relationship is considerably weaker for occupancy misreporting (Panel B). The effect is strong for appraisal overstatements (Panel C).

<<FIGURE 2 HERE>>

B. Regression Analysis

We now turn to a more formal framework. We estimate logit regressions where the dependent variable is the delinquency dummy and the independent variables of interest are the different misreporting indicators. To ensure that our misrepresentation variable is not simply capturing a correlation with some other aspect of loan riskiness, we control for the typical determinants of loan performance found in previous literature (Mayer et al. (2009)). In addition, we also include controls for complex mortgages, the original interest rate (for fixed and adjustable-rate mortgages (ARMs) separately), reported second liens, and reported occupancy status.[14] To allow for coefficients to be interpreted more easily, all continuous variables are standardized by subtracting their mean and dividing by their standard deviation. CBSA-Quarter fixed effects are included in all specifications (to ensure that our variables are not capturing some correlated aspect of regional home price movements). We also cluster standard errors by CBSA-Quarter. Results in Table III, Panel A, show the odds ratios and z-statistics (in parentheses) of the regressions when including *Unreported Second* in the set of

[14] A precise definition of each variable used in the regression is available in the Internet Appendix A.

explanatory variables. After controlling for the strict set of controls and fixed effects, we find that a first lien which has an unreported second lien is 97% more likely to become seriously delinquent than loans that were not misreported. The strong effect on loan performance of our indicator of second-lien misreporting is not driven by loans originated in California or Florida, since the odds ratio remains exactly the same when excluding these two states from the main sample. Panels B and C show that the effect of occupancy misreporting and appraisal overstatements on loan performance is lower than the effect of unreported second liens, though it is still important. Loans that misreport their occupancy status are 8% more likely to become delinquent than truthfully reported loans. As with second-lien misreporting, results are not driven by California or Florida. The effect of appraisal overstatements is also material. Loans which have inflated appraisals at the time of origination are 34% more likely to become delinquent. Results are not driven by loans originated in California or Florida. Finally, Panel D shows the results when using our aggregate indicator of misreporting. A misreported loan is associated with a higher likelihood of becoming delinquent of 51%. The effect of all the misreporting indicators is also economically and statistically significant when analyzing purchases and refinances separately. Finally, the results of the regressions in the common sample confirm that the misreporting indicator with most influence on performance is *Unreported Second*, followed by *Appraisal Overstatement*, followed by *Occupancy Misreporting*. Interestingly, the reported variables are quite important in predicting delinquencies, demonstrating that not all information was useless. Indeed, since the reported second lien indicator is the most important in terms of predicting default, it may explain why tinkering with this indicator would be valuable if one was to engage in misreporting.

In the Internet Appendix, Table IA.IV, we present OLS estimates of regressions like the ones discussed above but using *Direct Default* as the dependent variable. The effect of the misreporting indicators is also material. Loans that exhibit an unreported second lien, occupancy misreporting, or an overstated appraisal show a probability of going directly into foreclosure that is 52.3% (4.29%/8.2%), 11.8% (0.97%/8.2%), and 28.5% (2.34%/8.2%)

higher than the direct default mean rate of 8.2%, respectively. We find similar results when using foreclosures and early delinquencies as performance variables (as shown in Table IA.V).

<<TABLE III HERE>>

V. What is the Role of Borrowers, Appraisers, Lenders, and Underwriters in Misreporting?

We start by studying if lenders price the different measures of misrepresentation into the interest rates they charge borrowers. We then examine what we can learn regarding the role of lenders from credit score discontinuities, following a modified approach from Keys et al. (2009, 2010).

A. Was Misreporting Adequately Priced In by Lenders?

If lenders in fact take the features associated with the loans into account when they set interest rates, then it indicates that they recognize that the loans could have a higher level of risk. We regress loan interest rate at origination against the misreporting indicators and our strict set of loan level controls and fixed effects. Table IV shows the coefficient estimates for the variables of relevance (t-statistics in parentheses). Panel A shows that second-lien misreporting, which is the most prominent and impacts performance the most, is also associated with the largest increase in interest rates (14 bps). This seems to indicate that lenders knew about these second-lien loans. In fact, the higher interest rate charged is 4 bps larger than that of loans with reported second-lien loans (the difference is statistically significant, with an F-statistic of 25.37). Loans with occupancy misreporting (Panel B) show an interest rate increase of 5 bps on average. The interest rate is significantly lower than that of loans reported as investments or for second homes. Loans that exhibit appraisal overstatements seem to have a slightly higher interest rate at origination on average (7 bps). In summary, lenders seem to detect and internalize unreported second liens, and, to a lesser extent, appraisal overstatements. Occupancy misreporting seems to be instigated by

buyers, or at least originating lenders do not require extra compensation for the loans for owner-occupied purchases that may not be the borrower's primary residence.

<<TABLE IV HERE>>

B. Did Securitization Provide Incentives to Misreport?

If misrepresentation increases around a credit score used for securitization, then it seems likely that the originator is either intentionally or unintentionally facilitating borrowing with improper disclosure to obtain loans with the objective to securitize. If there is no jump in the amount of misrepresentation around the credit score threshold, then the misrepresentation is not a function of the originator screening process.

As background, we find results similar to Keys et al. (2010): low/no-documentation loans whose associated credit score is slightly over 620 are significantly more likely to be securitized.[15] Keys et al. (2010) analyzed data similar to ours and did not find evidence of credit scores being manipulated. They focus on delinquencies around the thresholds, whereas we center our analysis on the amount of potential misrepresentation around the thresholds.

We take the standard Regression Discontinuity Design (RDD) approach and normalize credit scores as follows:

$$C = Credit\ Score - Threshold, \tag{1}$$

where *Threshold* is 620 for low/no-documentation loans and 580 for full-documentation loans.

To distinguish credit scores that are over the threshold from credit scores that are below the threshold, we define:

$$D = \begin{cases} 1\ ,\ if\ C \geq 0 \\ 0\ ,\ otherwise. \end{cases} \tag{2}$$

Finally, a fourth order polynomial is fitted both above and below the credit scores thresh-

[15] Full-documentation loans whose associated credit score is slightly over 580 are also significantly more likely to be securitized, as shown in the Internet Appendix, Figure IA.5. For full-documentation loans we find the discontinuity at 600 shown in Keys et al. (2010), but at 580 the discontinuity in our sample is larger.

olds by means of the following specification:

$$Pct.\ Misreporting = \alpha + \beta D + \sum_{k=1}^{4} \gamma C^k + \sum_{k=1}^{4} \delta DC^k + \epsilon, \tag{3}$$

where *Pct. Misreporting* is a vector of the percentage of loans that exhibit potential misreporting for each credit score level (we run regressions for unreported second lien, appraisal overstatement, and occupancy misreporting separately).

Panel A of Figure 3 displays the results for unreported second liens. The amount of unreported second liens increases significantly in loans with credit scores of 620 with respect to loans with a credit score of 619, for both low/no-doc and full-doc loans. The percentage of unreported second liens increases by 2.5% for low/no-doc loans and 7.3% for full-doc loans. This suggests that this type of misreporting is a consequence of the originator's incentives to securitize. We find a small significant increase of 0.8% in the amount of occupancy misreporting only in low/no-documentation loans (Panel D). We do not find a significant increase in the amount of misreporting for appraisal overstatement (Table IA.VII and Figure IA.6). This suggests that these forms of misreporting are not directly related to the originator's motive to securitize.

<<FIGURE 3 HERE>>

C. Was Second-Lien Misreporting Facilitated Unintentionally or Intentionally?

The jumps in the probability of second-lien misreporting at the credit score thresholds of low/no-doc and full-doc loans suggest that originators facilitated misreporting either unintentionally (lax screening process) or intentionally (misrepresentation). In order to distinguish between the two possible explanations we further decompose our unreported second lien indicator into two components: (i) unreported second liens where the second lien was originated by the same lender that originated the first lien, and (ii) unreported second liens where the second lien was originated by a different lender than the one that originated the first lien. From the 10.2% second-lien misreporting we found in our sample, 67.6% (6.9%

/10.2%) consists of cases in which the same lender originated the first-second lien pair. The remaining are cases where the first and second lien had different originators. The fact that more than two-thirds of the second-lien misreporting occurs among loans originated by the same lender is surprising. Unless the bank had extremely poor record keeping, the issuer should know about the second lien. The second-lien misreporting seems likely due to intentional misreporting on the part of the originators or from MBS underwriters who realize the loan has a second lien but do not report this information to investors.

We repeat the RDD Analysis discussed above for the two types of second-lien misreporting. Panels B and C of Figure 3 show the discontinuity results for the same originator and different originators. Panel B shows that when the same originator is on both the first and second lien, there is a misreporting jump of 5.5% (significant at the 1% level) at the credit score threshold of 580. For low/no-doc, the amount of misreporting increases rapidly over the credit score threshold of 620, reaching the same levels present in full-docs over 580. Nevertheless, because of the scattered loan pools between the 600 and 620 marks, there is no statistically detectable jump in the case of low/no-docs. Likewise, Panel C shows that when different originators issue the first and second liens, there is also a significant jump in second-lien misreporting (1.7% for low/no-docs and 1.9% for full-docs). These results reaffirm that second-lien misreporting, at least in part, occurs due to the incentives of the lender to securitize the loan.

Nevertheless, it is not clear if the misreporting is due to the loan originator or bank underwriter. It could be that the originators intended to securitize and reported the second lien information properly to the bank underwriters, who later did not report it. To understand if the misreporting was a function of the loan originators or MBS underwriters, we regress our misreporting indicators on loan originator and underwriter fixed effects. If loan originators contributed more to certain type of misreporting, then the loan originator fixed effects should be more important in explaining the misreporting. If underwriting practices contributed more, then underwriter fixed effects should explain more of the misreporting

variation.

In Table V we also include CBSA-Quarter fixed effects and other controls. For second-lien misreporting, loan originator fixed effects explain a larger proportion of the second-lien misreporting variation than underwriter fixed effects. The regression of all controls with CBSA-Quarter fixed effects (but neither originator nor underwriter fixed effects) yields an R^2 of 0.10. Adding originator but not underwriter fixed effects yields an R^2 of 0.152, compared to only 0.129 with only adding underwriter fixed effects. This suggests that second-lien misreporting, at least initially, seems to be more aligned with originator practices rather than with underwriter practices. Nevertheless, since underwriter fixed effects explain additional variation beyond originator fixed effects (R^2 of 0.168), second-lien reporting may also be a decision of the underwriter. For appraisal overstatement and owner occupancy misreporting, Table V shows that the originator and underwriter fixed effects do not explain much of the misreporting. These types of misreporting did not vary widely across originators and underwriters.

<<TABLE V HERE>>

D. Why are Appraisals Overstated?

Our analysis has shown that it is much more common for appraisals to be substantially above the AVM model rather than the opposite, and that these appraisal overstatements are strongly related to future loan performance. This pattern could be prevalent because: (i) appraisers do their best but make random mistakes in their appraisals, or (ii) appraisers target the expectations of loan officers, and thus their appraisals may be upward biased. To examine which of these two possibilities is more prevalent, we examine refinances. With purchases as long as the transaction is arms-length, the buyer has an incentive to purchase at a low price. But with refinances, the price of the house rest solely on the appraisal. If appraisers are in general trying to please loan officers, then we should expect to see more inflated appraisals for refinances. Among refinances, appraisal inflation might be largest in

the cash-out refinances where the buyer's goal is not just to repay the previous debts on the property, but also to maximize the loan value taken. The loan officer may also have an incentive to maximize the loan size, as his or her commission is proportional to the dollar value of the loan.

As shown in Table II, appraisal overstatements are 13.2% for purchases and 20.5% for refinances, a 55.3% relative increase. In Panel A of Table VI, we estimate logit regressions for the frequency of appraisal overstatement and confirm that the higher levels of overstatements for refinances are not driven by loan characteristics. Appraisals in refinances are 74% more likely to be overstated than purchases. Additionally, we find that cash-out loan appraisals are more likely to be overstated than term refinance loan appraisals (odds ratio of 1.81 compared to 1.57, different at the 1% level).

<<TABLE VI HERE>>

We delve deeper into the two possible explanations for the predictive ability of appraisal overstatements for refinances. With refinances, we exclude loans where there are second liens and focus on loans that occur at LTV increments of five. We might see loans clustering at five-unit increments for two reasons. Consistent with explanation (i) above, the loan officer asks for an unbiased appraisal and then loans up to the value of the appraisal which will cluster at increments of five. Consistent with explanation (ii), the loan officer decides how much the value of the loan needs to be and then tells the appraiser what valuation to target. Consistent with both explanations, 49% of appraisals in refinances with no second lien are at or within 0.5 units of a five-unit increment. This is shown by the green bars in Figure 4.[16] For term refinances, the home owner is only paying back previous debt. To the extent that the home price has changed, there is little reason under the first explanation to think that the values of the house should cluster exactly at five units increments. Nevertheless, we find that amount of loans occurring at five-unit increments is still high for term repurchases

[16] So that only the amount of the first-lien mortgage is relevant, we remove loans with second liens (either reported or unreported). Interestingly, most of this activity happens exactly at the five unit loan-to-value number.

(33%).

We examine the amount of appraisal overstatement at the five-unit increments. If appraisal overstatements are random errors (explanation (i)), overstatements should not be more or less likely at five-unit increments. The targeting explanation (ii) suggests that appraisals are more overstated at five-unit increments since the appraiser would be told to deliver a high appraisal. As shown by the blue circles in Panel A of Figure 4, appraisal overstatements are consistently higher on every five-unit increment by 1% to 2%. Panel B shows that loans on the five-unit increments also default at a much higher rate. Both of these findings are consistent with appraisers following instructions from loan officers to target high valuations to meet a certain loan-to-value. Finally, in Panel C we examine the relationship between the level of appraisal overstatement and default for those loans at five-unit increments and those not at the five-unit increments. Delinquency rates are considerably higher across loan-to-value ratios for loans that fall on five-unit increments. Additionally, the delinquency rates rise rapidly from 0% to 5% overstated. Even loans with a 5% appraisal overstatement default at a much higher rate, and this increase is more apparent for loans at five-unit increments.

<<FIGURE 4 HERE>>

If appraisals on five-unit increments are driven by targeting values from loan officers, the overstatement indicator at five-unit increments should be more important for predicting default. We go back to our logit delinquency regressions. Panel B of Table VI shows that for loans at five-unit increments, appraisal overstatements lead to a further increase in the likelihood of delinquency (the odds ratio of appraisal overstatements at five-unit LTV increments is 1.25 compared to 1.16 not at the increments).

Our analysis presents strong evidence that our appraisal overstatement variable captures a significant aspect of misreporting by appraisers and the loan officers. It also suggests that the success of the appraisal overstatement indicator at capturing delinquency is strongly related to appraisal targeting and not just to random errors by appraisers. Nevertheless,

random appraisal errors likely play a role as well.

E. What Role did MBS Underwriters Play in Misreporting?

So far, we have focused on the possible role of borrowers, appraisers, and loan originators, but not on that of MBS underwriters. Owner occupancy percentages, loan-to-value ratios, and second lien percentages are typically reported in MBS prospectuses. Detailed loan-level data, including the second-lien information we use, typically originates from raw servicer/trustee loan-level data tapes, but is also similarly displayed in many free writing prospectuses. From the testimony of Vicki Beal of Clayton Holdings before the Financial Crisis Inquiry Commission, we know that firms like Clayton Holdings performed due diligence of MBS pools and provided this information to most MBS underwriters. For MBS pools, a sample of approximately 3-10% of the underlying loans was picked and sent to a firm like Clayton Holdings for detailed loan-level monitoring.[17] Interestingly, for sample deals provided in the congressional inquiries, Clayton found that about 28% of loans did not meet credit and compliance guidelines. Yet, the banks would typically conclude that the problem was limited to only the sample and replace only part of the troubled loans.[18] Additionally, if assertions are true that monitoring firms had lax practices, then the true extent of misrepresentation could be substantially higher.

Silent second, owner occupancy, and appraisal verification are all types of services listed by Clayton Holdings and long used by monitoring services. Nevertheless, the publicly released Clayton information by the FCIC is not detailed enough to see what exact standards Clayton used for MBS monitoring. It is worthwhile to note that although the originator fixed effects explain more of the variation of the unreported seconds indicator in Table V,

[17]CoreLogic, Fidelity Information Services, 406 Partners, Allonhill, American Mortgage Consultants, Opus Capital Markets Consultants, and RR Donnelly are listed by Ms. Beal as competitors. Yet, Clayton Holdings held up to 70% of the market (Nasiripour (2011)).

[18]Underwriters may have picked random samples or samples with more favorable or adverse characteristics. There was then a waiving process where even for the sample loans, 'exceptions' would be made. So, for example, for a MBS with 10% loans sampled and 28% of these loans found not to have been up to underwriting standards, only 1.6% might be replaced. According to Vicki Beal, underwriting firms used this information to negotiate better prices on the loan pools they were purchasing.

the underwriter fixed effects still explain a considerable amount of variation. This indicates that MBS underwriting disclosure practices mattered even beyond the originators that they chose to underwrite.

VI. Is Originator Performance Related to Misrepresentation Indicators?

Given the substantial amount of misreporting in non-agency MBS pools and its negative effect on loan performance, a natural question that arises is whether misreporting was widespread or concentrated in a few originators.

A. Performance and Second-Lien Misreporting

As showed in the previous section, originators seemed to be aware of second-lien misreporting. We start by evaluating mortgage lenders in terms of the performance of the loans that they originated, and then we relate performance to the amount of second-lien misreporting.

We analyze the loans issued by the largest 25 mortgage originators in the main sample.[19] We estimate an OLS regression where the dependent variable is the delinquency dummy and the explanatory variables are the loan-level controls used in Table III,but we do not include the misrepresentation indicators. CBSA-quarter and originator fixed effects are also included. The variable of interest is the originator fixed effect estimate. This estimate captures the positive/negative excess delinquency rate experienced by an originator after controlling for observed risk, relative to an originator of reference. We interpret originators with the highest originator fixed effects estimates as having a worse origination process than their counterparts with the lowest estimates. Panel A of Figure 5 plots the originator fixed effect estimates on the vertical axis and the percentage of misreporting exhibited by each lender on the horizontal axis. Several things can be noticed. First, there is significant

[19]We have originator names for 88.3% of the loans in the main sample, 81.6% of which were issued by the largest 25 originators. A complete list with the names of these originators and their abbreviations is provided in the Internet Appendix, Table IA.XIII.

variation in the quality (performance) across lenders. Fixed effects range from -0.101 to 0.078, which implies that loans originated by the best performing originator default 17.9% less on average than the loans originated by the worst performing originator. Second, there is also significant variation in the amount of second-lien misreporting across lenders with misreporting ranging from 0.56% to 40.2%.[20] This suggests that some lenders played a more important role in facilitating misrepresentation than others. Third, there is a positive relationship between performance and misreporting across lenders. The fixed effect estimate and the amount of second-lien misreporting yield a positive correlation of 0.79 (significant at the 1% level). Panel A of Figure 5 indicates that poor originator performance is not just random, but strongly correlated with the prevalence of the average amount of potential misrepresentation within the loan originator.

<<FIGURE 5 HERE>>

B. Do the Misrepresentation Indicators Capture the Full Extent of Misrepresentation?

We examine the same relationship as in the previous subsection, but we remove the original effects of our misrepresentation indicators. If the misrepresentation indicators are capturing the full extent of mortgage misrepresentation, then there should be no relationship between originator fixed effects (after controlling for mortgage misrepresentation) and the extent of second-lien misreporting at each originator. Panel B of Figure 5 plots originator fixed effects, except that this time, we add the three misreporting variables as additional controls to the regression in Subsection A. The correlation between the originator fixed effect estimates and the amount of second-lien misreporting weakens only slightly; there is still a positive correlation of 68%. The amount of misrepresented loans of the lender remains strongly related to lender performance, even after controlling for the loans flagged

[20]Interestingly, WMC (a subsidiary of GE Capital) which has the highest second-lien misreporting rate in our sample, has faced serious accusations that misrepresentation and document falsification was rampant and that several whistleblowers were completely ignored and sidelined. As of January 2012, the LA Times reports that they are under criminal investigation by the FBI and the US Department of Justice for falsifying paperwork and other questionable loan practices (Hudson and Reckard (2012)).

as potentially misrepresented. Lenders who have large amounts of misrepresented loans have abnormal negative performance due to either more misrepresented loans or some other aspect of poor originating practices correlated with the extent of mortgage misrepresentation but not captured in our detailed loan-level data controls.

C. Does Misreporting Around Securitization Thresholds Vary Across Lenders?

We seek to understand how mortgage originator performance varies with mortgage misrepresentation indicators around securitization thresholds. We define the best (worst) performers as the tercile of originators with the smallest (largest) originator fixed effects based on the specification described in Subsection A. We repeat the RDD analysis presented in Section V but only consider these two subsets of lenders. Figure 6 shows that for loans originated by bad lenders, the amount of second-lien misreporting increases 3.2% at the credit score thresholds of 620 for low/no-documentation loans and increases 14.3% at 580 for full-documentation loans. These values are much larger than the jumps for the best lenders, which show a small negative jump of 1.5% in the case of low/no-docs and a jump of 4.2% in the case of full-docs.

For both poor and best performing lenders, more than two-thirds of loans are originated by the same lender. Hence, both good and bad performing lenders should be aware of unreported second liens, though poor performers seem to be intentionally facilitating second-lien misreporting to a higher degree than good performers.

<<FIGURE 6 HERE>>

D. Is Misreporting Persistent?

We examine if misreporting within lenders is persistent and if this had implications for the broader market. For this means, we split the sample in two: loans originated during 2002-2005 and loans originated during 2006-2007. Once again we estimate the specification used in Subsection A for each of these subsamples separately. Figure 7 shows that second-lien

misreporting is persistent within the lenders for the two periods. The correlation between periods is 0.64.[21] Practices of securitizing misrepresented loans were pervasive and persistent across originators for quite some time.

<<FIGURE 7 HERE>>

Previously, in Figure 1, we documented a massive run-up and run-down in unreported second liens. It is natural to ask how the persistent behavior of second-lien misreporting can be reconciled with such a pattern. To analyze this question, in Figure 8 we examine originators in the top and bottom terciles of second-lien misreporting in 2006. The bottom tercile with the highest amount of misreporting was issuing high levels of poor loans back to the beginning of our sample in 2002. Even though the proportion of origination volume from issuers with high second-lien misreporting was small in the early years, it grows and peaks in 2005-2006. In the third quarter of 2007, issuance volume dropped dramatically back to 2002 levels. Thus, the overall dramatic rise and fall in second-lien misreporting seems to be due to issuers with high levels of misreporting capturing a larger share of the market and ultimately falling due to financial trouble from poor performing issuances in prior years. This is again consistent with persistent misreporting but wide variation in misreporting across issuers. The findings also indicate that poor loan performance is not an accident but a function of lender practices.

<<FIGURE 8 HERE>>

VII. Did Misreporting Contribute to MBS Performance?

Although misreporting is related to poor individual loan performance, it is not clear how relevant loan losses are for aggregate MBS pool performance. We analyze a cross-section of 1,887 MBS and focus on quantifying the effects of mortgage misreporting on potential future losses of the MBS.[22]

[21] Figure IA.10 shows that occupancy misreporting and appraisal overstatements are also persistent with correlations of 0.52 and 0.69 between periods.

[22] We describe the sample selection process and how we compute the MBS projected losses in the Internet Appendix B.

Panel A of Table VII summarizes the results of OLS regressions on projected MBS losses. The coefficient of 0.08 on the second-lien misreporting variable implies that if potentially misrepresented loans in the pool increase by one standard deviation in a deal with projected losses equal to the sample mean of 15.9%, the amount of losses in the MBS increases by 1.18% ((exp(0.08×0.127)-1)×(1+0.159)) on average. This is equivalent to a 7.4% increase in projected losses relative to the mean. This effect is economically important given that the average size of the equity tranche is around 1.2% (see Begley and Purnanandam (2013)). Likewise, the effects of occupancy misreporting, appraisal overstatements, and aggregated misreporting on MBS losses are also material. An increase of one standard deviation in misreporting increases MBS losses in 6.6% for occupancy misreporting, 8.1% for appraisal overstatements, and 11.5% for aggregate misreporting (relative to the mean). Additionally, Panel B of Table VII (using the coefficient estimates of the regression that uses aggregate misreporting) shows that the impact of misreporting on MBS performance is even larger than variables that are commonly thought to be of main importance such as LTV and percent of low/no-doc loans. In summary, misreporting seems to have a significantly large economic effect on projected MBS losses. Piskorski, Seru, and Witkin (2013) show that the effects of misreporting were not reflected in MBS pricing or subordination. Thus it seems likely that investors were adversely affected by misreporting.

<<TABLE VII HERE>>

Finally, we examine differential reporting by underwriters. The top graph in each panel of Figure 9 shows misreporting by underwriters. Panel A shows there are substantial differences in the amount of second-lien misreporting across underwriters.[23]

It could be that the large cross-sectional differences in misreporting levels shown by underwriters is due to the securitization of loans issued by originators with high levels of misreporting. To examine this, we construct a benchmark for misreporting based on the average amount of misreporting of each originator. We gauge abnormal misreporting by an

[23]The list of underwriters and corresponding abbreviations are presented in Table IA.XIV.

underwriter as the actual misreporting minus the benchmarked misreporting. Underwriters with high second-lien misreporting tend to show more misreporting than the benchmark. This indicates that the cross-sectional differences in misreporting across underwriters are due in part to differences in the quality of their second-lien practices. For owner occupancy and appraisal overstatements, there are only small differences across underwriters after controlling for the average for originators.

From the 18 largest players in the securitized market, the highest aggregated misreporting is by Barclays and JP Morgan (41.5% and 41%, respectively). Consistent with our findings, as part of a recent 13 billion dollar settlement with the Department of Justice, JP Morgan admitted mortgage misrepresentations on MBS they issued.[24] Nevertheless, misreporting is high for all underwriters in general; the least amount of misreporting is 23.6% for Washington Mutual. It is also important to remember that our analysis only covers three types of mortgage misrepresentation, and only uses very conservative estimates so there is likely much more misrepresentation than we can detect. Overall, the MBS analysis shows that the effects of misrepresented loans are not just abstract concepts, but associated with substantial differences in misreporting and losses to actual MBS pools.

<<FIGURE 9 HERE>>

VIII. Conclusion

Using a large sample of non-agency securitized loans originated from 2002 to 2007, we find sizeable amounts of potential mortgage misrepresentation in the form of unreported second liens, owner occupancy misreporting, and inflated appraisals. These apparent misrepresentation patterns are surprisingly similar for full and low/no-documentation loans. Loans with a misrepresentation indicator are 51% more likely to become seriously delinquent; these indicators, which can be measured at issuance, are widely informative regarding future loan performance. Since lenders do not charge higher interest rates for loans with owner occupancy

[24]The settlement explicitly admits mortgage misreporting from securities issued by JP Morgan and not just from Bear Sterns and Washington Mutual.

misreporting and the amount does not vary around securitization thresholds, this seems to be a misrepresentation largely facilitated by borrowers. However, higher interest rates for loans with unreported second-lien loans and loans with inflated appraisal values suggest that lenders were, to some extent, aware of this risk. In more than two-thirds of the cases where a first-lien loan has an unreported second-lien loan, both loans were issued simultaneously by the same originator. Second-lien misreporting among low and full-documentation loans is considerably greater around credit score thresholds, suggesting that second-lien misreporting, at least in part, occurs because the originator intended to securitize the loans. Appraisal overstatements are much greater around loan-to-value thresholds, and these thresholds are associated with higher default rates, indicating that the appraisers were commonly targeting valuations from loan officers.

Owner occupancy and appraisal overstatement misreporting does not vary widely across originators and underwriters. The importance of both originator and (to a lesser extent) underwriter fixed effects in explaining cross-sectional differences in whether a second lien is misreported suggests that the misreporting was a function of both the originator and the underwriter. Even after removing the component of performance explained by our misreporting indicators, originator loan performance is strongly related to originator-level second lien misreporting. This suggests that our misreporting indicators are not capturing the full extent of misreporting or other poor practices. The percentage of pool level misreporting at MBS issuance is one of the most economically important predictors of future MBS losses, suggesting that at least MBS losses could have been anticipated.

In the sense that accurate weights and measures are essential for trade, an accurate description of an asset seems to be a minimum condition for a sustainable market. Surprisingly, these basic conditions do not seem to have been met on a wide-scale basis. Our results are generally consistent with the originate-to-distribute explanation, but we are not attempting to provide a comprehensive explanation of causes of the mortgage crisis, nor do we focus on what was known by investors. Beyond lax screening, our findings indicate that the problems

with non-agency mortgage-backed securitization ran even deeper than incomplete screening, a decline in MBS standards, and a lack of available information. While originators not considering ultimate loan performance is consistent with lax screening, our results suggest that originators seemingly possessed information to realize that their loans were considerably worse than represented to investors. Large-scale misreporting even for full-documentation loans indicates that securitization cannot be fixed simply by requiring more documentation.

As one considers the future of securitization, investors may need additional recourse and guarantees that someone will more consistently stand behind the stated representations of the underlying assets. While the securitization market has re-emerged at a vibrant pace (CMBS, ABS, and some non-agency MBS), one must wonder if the (seemingly minor) changes made within the market are sufficient given the major structural problems uncovered in the past and the possible negative banking externalities that the market has been shown to generate. Perhaps additional research will continue to enhance the understanding of this complex market and lead to more transparency.

References

Agarwal, Sumit, and Itzhak Ben-David, 2013, Do Loan Officers' Incentives Lead to Lax Lending Standards?, Working paper.

Agarwal, Sumit, Itzhak Ben-David, and V. Yao, 2014, Collateral Valuation and Institutional Pressures: Evidence from the Residential Real-Estate Market, *Management Science,* forthcoming.

Akerlof, George A, 1970, The Market for "Lemons": Quality Uncertainty and the Market Mechanism, *Quarterly Journal of Economics* 26, 488–500.

Amromin, Gene, Jennifer Huang, Clemens Sialm, and Edward Zhong, 2013, Complex Mortgages, Working paper.

Ashcraft, Adam, Paul Goldsmith-Pinkham, and James Vickery, 2010, MBS Ratings and the Mortgage Credit Boom, FRB of New York Staff Report.

Becker, Gary S., 1968, Crime and Punishment: An Economic Approach, *Journal of Political Economy* 76, 169–217.

Begley, Taylor, and Amiyatosh Purnanandam, 2013, Design of Financial Securities: Empirical Evidence from Private-label RMBS Deals, Working paper.

Ben-David, Itzhak, 2011, Financial Constraints and Inflated Home Prices during the Real Estate Boom, *American Economic Journal: Applied Economics* 3, 55–87.

Bernardo, Antonio E., and Bradford Cornell, 1997, The Valuation of Complex Derivatives by Major Investment Firms: Empirical Evidence, *Journal of Finance* 52, 785–798.

Bolton, Patrick, Xavier Freixas, and Joel Shapiro, 2007, Conflicts of interest, information provision, and competition in the financial services industry, *Journal of Financial Economics* 85, 297–330.

Bolton, Patrick, Xavier Freixas, and Joel Shapiro, 2012, The Credit Ratings Game, *Journal of Finance* 67, 85–111.

Booth, James R., and Richard L. Smith, 1986, Capital raising, underwriting and the certification hypothesis, *Journal of Financial Economics* 15, 261–281.

Carrillo, Paul E, 2011, Testing for Fraud in the Residential Mortgage Market: How Much Did Early-Payment-Defaults Overpay for Housing?, *Journal of Real Estate Finance and Economics* 1–29.

Chemmanur, Thomas J., and Paolo Fulghieri, 1994, Reputation, Renegotiation, and the Choice between Bank Loans and Publicly Traded Debt, *Review of Financial Studies* 7, 475–506.

Coval, Joshua D, Jakub W Jurek, and Erik Stafford, 2009, Economic Catastrophe Bonds, *American Economic Review* 99, 628–666.

DeMarzo, Peter M, 2005, The Pooling and Tranching of Securities: A Model of Informed Intermediation, *Review of Financial Studies* 18, 1–35.

Demiroglu, Cem, and Christopher James, 2012, How Important is Having Skin in the Game? Originator-Sponsor Affiliation and Losses on Mortgage-backed Securities, *Review of Financial Studies* 25, 3217–3258.

Demyanyk, Yuliya, and Otto Van Hemert, 2011, Understanding the Subprime Mortgage Crisis, *Review of Financial Studies* 24, 1848–1880.

Ehrlich, Isaac, 1973, Participation in Illegitimate Activities: A Theoretical and Empirical Investigation, *Journal of Political Economy* 81, 521–565.

Faltin-Traeger, Oliver, and Christopher Mayer, 2012, Lemons and CDOs: Why Did So Many Lenders Issue Poorly Performing CDOs?, Working paper.

Federal Bureau of Investigation, 2011, Financial Crimes Report 2010-2011.

Financial Crisis Inquiry Commission, 2010, The Impact of the Financial Crisis-Sacramento, Session 3.

Financial Crisis Inquiry Commission, 2011, The financial crisis inquiry report.

Foote, Christopher L., Kristopher S. Gerardi, and Paul S. Willen, 2012, Why Did So Many People Make So Many Ex Post Bad Decisions? The Causes of the Foreclosure Crisis, Working paper.

Fulghieri, Paolo, Gnter Strobl, and Han Xia, 2014, The Economics of Solicited and Unsolicited Credit Ratings, *Review of Financial Studies* 27, 484–518.

Garmaise, Mark J., 2013, Borrower Misrepresentation and Loan Performance, *Journal of Finance,* forthcoming.

Goodman, Laurie S., Shumin Li, Douglas J. Lucas, Thomas A. Zimmerman, and Frank J. Fabozzi, 2008, *Subprime Mortgage Credit Derivatives* (John Wiley & Sons).

Griffin, John, Richard Lowery, and Alessio Saretto, 2014, Complex Securities and Underwriter Reputation: Do Reputable Underwriters Produce Better Securities?, *Review of Financial Studies,* forthcoming.

Griffin, John, Jordan Nickerson, and Dragon Tang, 2013, Rating Shopping or Catering? An Examination of the Response to Competitive Pressure for CDO Credit Ratings, *Review of Financial Studies* 26, 2270–2310.

Griffin, John M, and Dragon Yongjun Tang, 2011, Did Credit Rating Agencies Make Unbiased Assumptions on CDOs?, *American Economic Review* 101, 125–130.

Griffin, John M, and Dragon Yongjun Tang, 2012, Did Subjectivity Play a Role in CDO Credit Ratings?, *Journal of Finance* 67, 1293–1328.

Grossman, Sanford J., and Oliver D. Hart, 1980, Disclosure Laws and Takeover Bids, *Journal of Finance* 35, 323–334.

Haughwout, Andrew, Donghoon Lee, Joseph Tracy, and Wilbert Van der Klaauw, 2011, Real Estate Investors, the Leverage Cycle, and the Housing Market Crisis, FRB of New York Staff Report.

He, Jie (Jack), Jun (QJ) Qian, and Philip E Strahan, 2012, Are All Ratings Created Equal? The Impact of Issuer Size on the Pricing of Mortgage-Backed Securities, *Journal of Finance* 67, 2097–2137.

Hudson, Michael, and E. Scott Reckard, 2012, GE lending unit said to be target of U.S. probe, *Los Angeles Times*, January 20.

Jiang, Wei, Ashlyn Nelson, and Edward Vytlacil, 2014, Securitization and Loan Performance: A Contrast of Ex Ante and Ex Post Relations in the Mortgage Market, *Review of Financial Studies,* forthcoming.

Jiang, Wei, Ashlyn Aiko Nelson, and Edward Vytlacil, 2013, Liar's Loan? Effects of Origination Channel and Information Falsification on Mortgage Delinquency, *Review of Economics and Statistics,* forthcoming.

Keys, Benjamin J, Tanmoy Mukherjee, Amit Seru, and Vikrant Vig, 2009, Financial regulation and securitization: Evidence from subprime loans, *Journal of Monetary Economics* 56, 700–720.

Keys, Benjamin J, Tanmoy Mukherjee, Amit Seru, and Vikrant Vig, 2010, Did Securitization Lead to Lax Screening? Evidence from Subprime Loans, *Quarterly Journal of Economics* 125, 307–362.

Keys, Benjamin J, Amit Seru, and Vikrant Vig, 2012, Lender Screening and the Role of Se-

curitization: Evidence from Prime and Subprime Mortgage Markets, *Review of Financial Studies* 25, 2071–2108.

Lee, David S., and Thomas Lemieux, 2010, Regression Discontinuity Designs in Economics, *Journal of Economic Literature* 48, 281–355.

Lee, Donghoon, Christopher J Mayer, and Joseph Tracy, 2012, A New Look at Second Liens, NBER Working paper 18269.

Mathis, Jerome, James McAndrews, and Jean-Charles Rochet, 2009, Rating the raters: Are reputation concerns powerful enough to discipline rating agencies?, *Journal of Monetary Economics* 56, 657–674.

Mayer, Christopher, Karen Pence, and Shane M Sherlund, 2009, The Rise in Mortgage Defaults, *Journal of Economic Perspectives* 23, 27–50.

McCrary, Justin, 2008, Manipulation of the running variable in the regression discontinuity design: A density test, *Journal of Econometrics* 142, 698–714.

Nasiripour, Shahien, 2011, Pandering To Clients, Firm Distances Itself From Wall Street Whistleblower's Testimony, *The Huffington Post*, May 25.

Piskorski, Tomasz, Amit Seru, and James Witkin, 2013, Asset Quality Misrepresentation by Financial Intermediaries: Evidence from RMBS Market, Working paper.

Povel, Paul, Rajdeep Singh, and Andrew Winton, 2007, Booms, Busts, and Fraud, *Review of Financial Studies* 20, 1219–1254.

Purnanandam, Amiyatosh, 2011, Originate-to-distribute Model and the Subprime Mortgage Crisis, *Review of Financial Studies* 24, 1881–1915.

Riddiough, Timothy J, 1997, Optimal Design and Governance of Asset-Backed Securities, *Journal of Financial Intermediation* 6, 121–152.

Sangiorgi, Francesco, and Chester Spatt, 2013, Opacity, Credit Rating Shopping and Bias, Working paper.

Tables and Figures

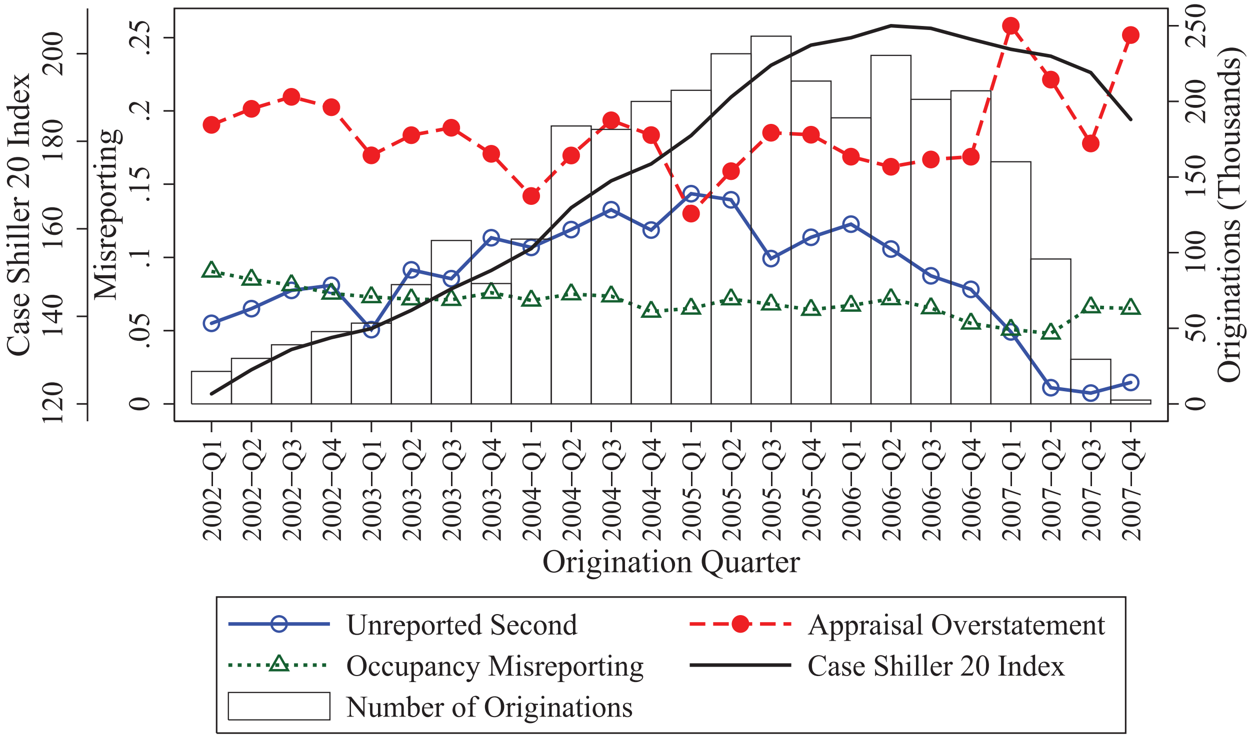


Figure 1. Misreporting Indicators By Quarter. This figure shows the evolution of the different misreporting indicators by quarter. The indicators *Unreported Second*, *Occupancy Misreporting*, and *Appraisal Overstatement* are defined in Table II. The bars represent the number of first-lien loan originations in each quarter, and the continuous black line shows the evolution of house prices (as captured by the Case Shiller 20 Index). The main sample consists of ABSNet-DataQuick matched securitized first-lien loans used for the purchase or refinance of a home with an initial loan amount over $30k and a loan-to-value ratio (LTV) lower or equal than 103%. We drop loans associated to the largest 1% of the transactions in each state, loans that are reported as being for homes of more than one unit, and loans that belong to MBS deals in which all mortgages are recorded to have an LTV equal to their combined LTV.

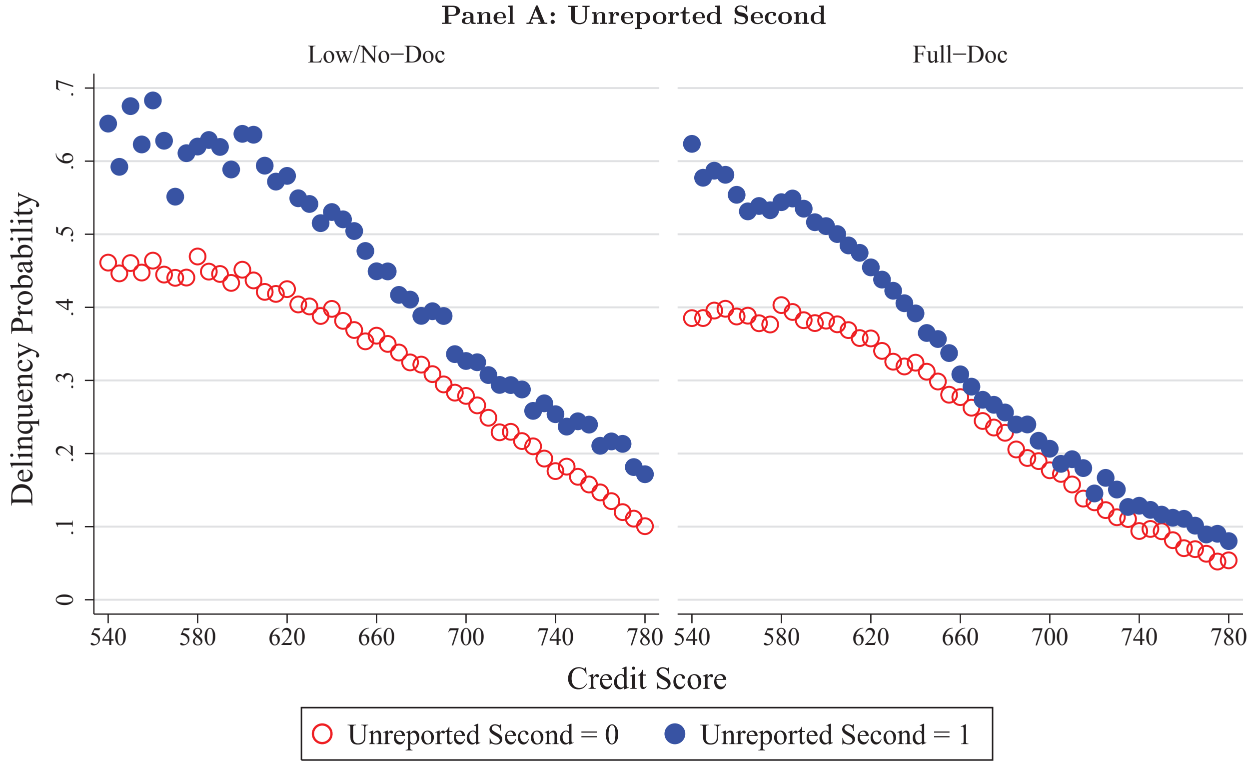
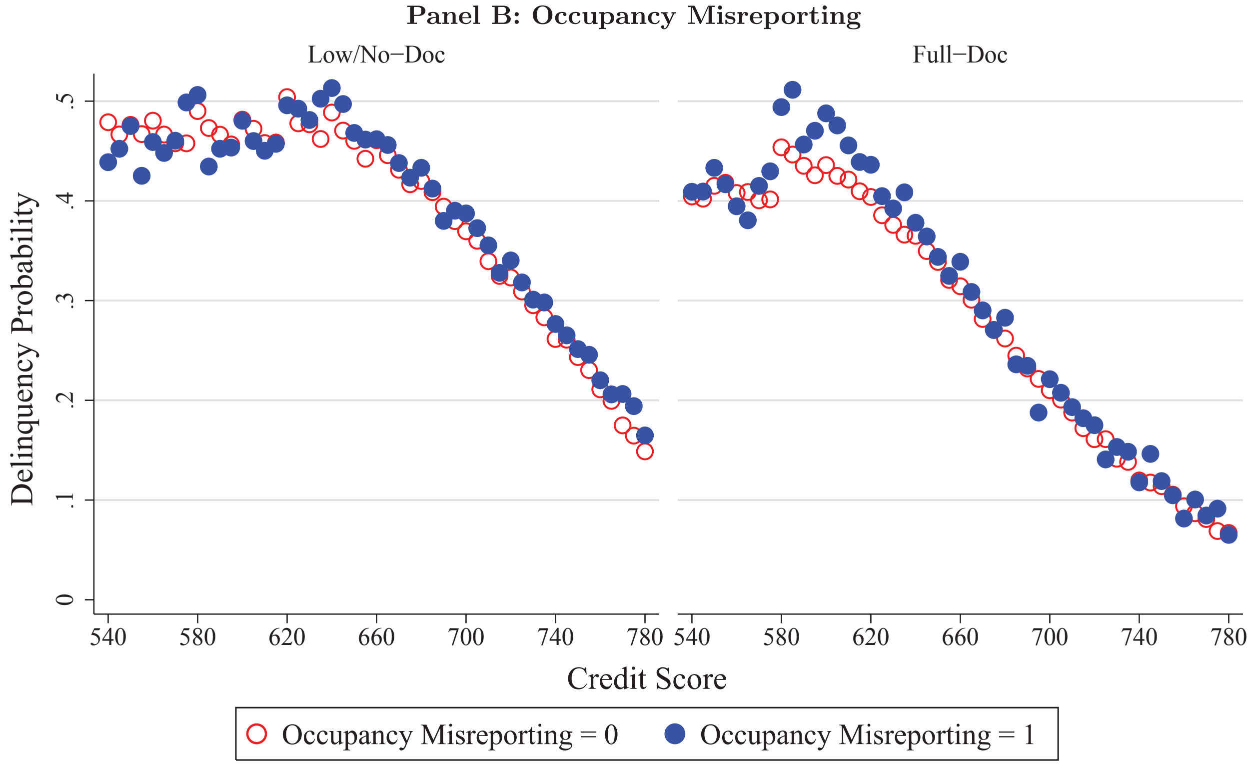
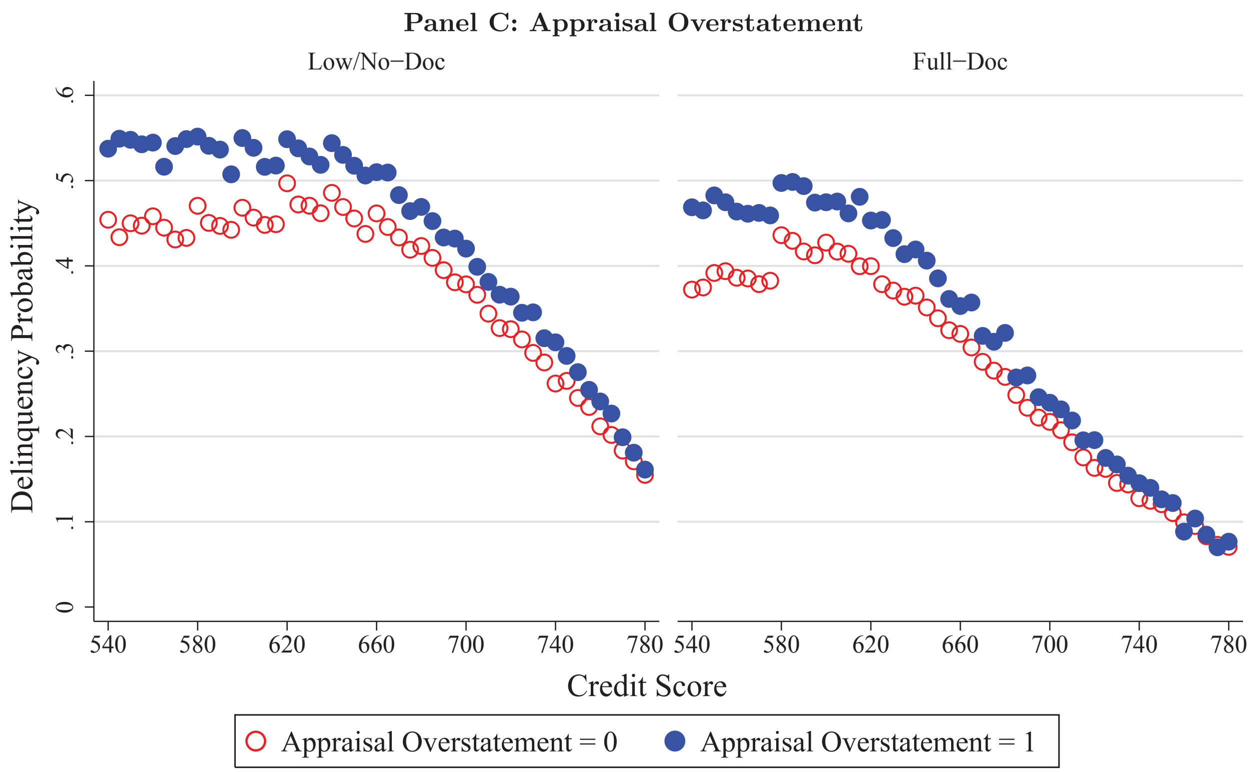


Figure 2. Probability of Delinquency By Credit Score Bin. This figure compares the probability of delinquency by credit score bin of loans that exhibit misreporting (blue circles) and loans that do not exhibit misreporting (hollow red circles). Each credit score bin has a size of five units. We construct three potential mortgage misreporting indicators to capture unreported second liens, occupancy status misreporting, and appraisal overstatements (the indicators are defined in Table II). The comparisons based on unreported second liens, occupancy misreporting, and appraisal overstatements are shown in Panels A, B, and C, respectively.


Panel B: Occupancy Misreporting
Low/No−Doc
Full−Doc
Delinquency Probability
0
.1
.2
.3
.4
.5
540
580
620
660
700
740
780
Credit Score
Occupancy Misreporting = 0
Occupancy Misreporting = 1


Panel C: Appraisal Overstatement
Low/No−Doc
Full−Doc
Delinquency Probability
0
.1
.2
.3
.4
.5
.6
540
580
620
660
700
740
780
Credit Score
Appraisal Overstatement = 0
Appraisal Overstatement = 1

Panel A: Unreported Second (All Loans)

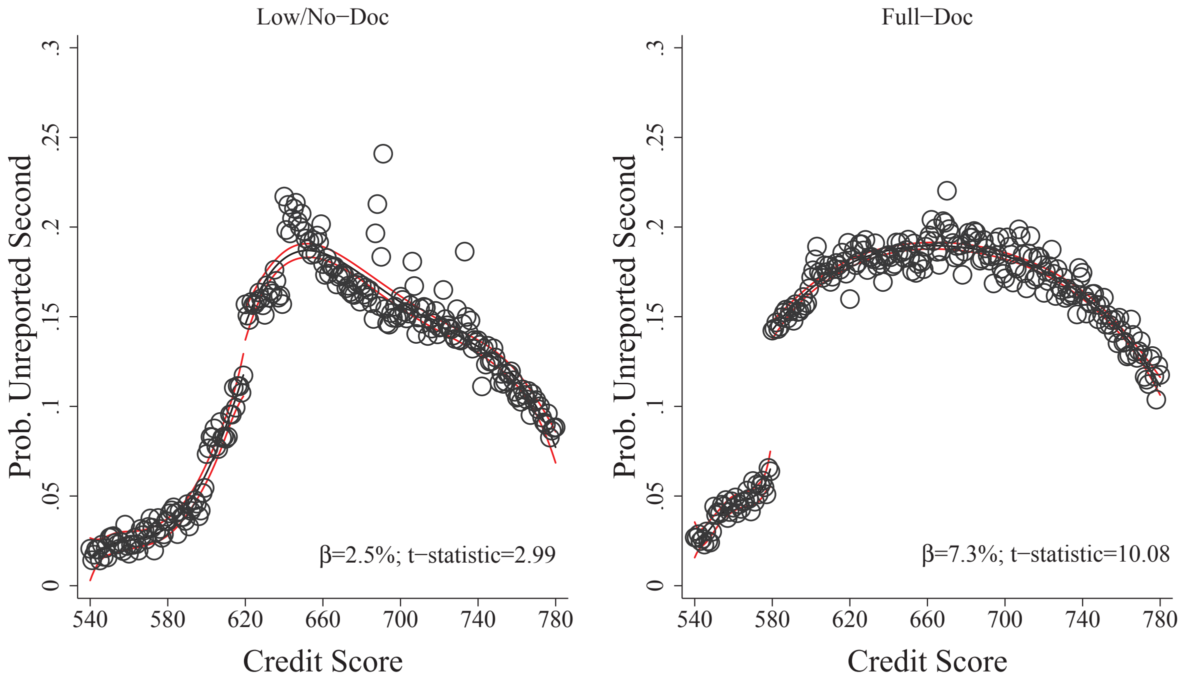
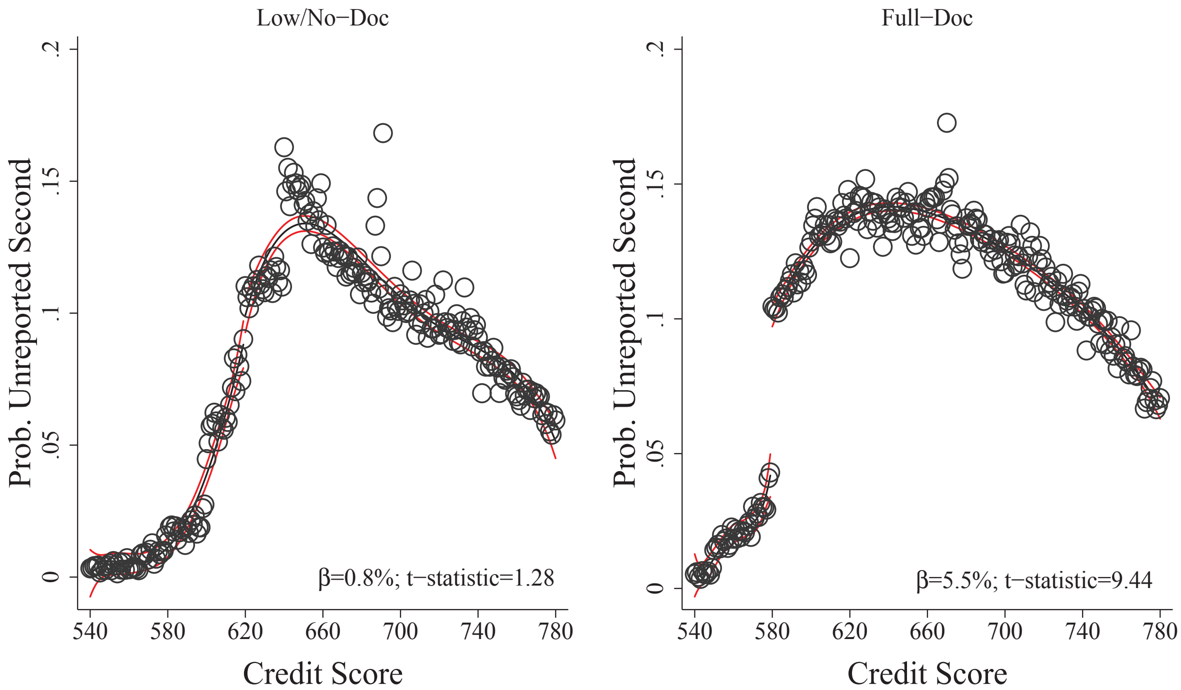
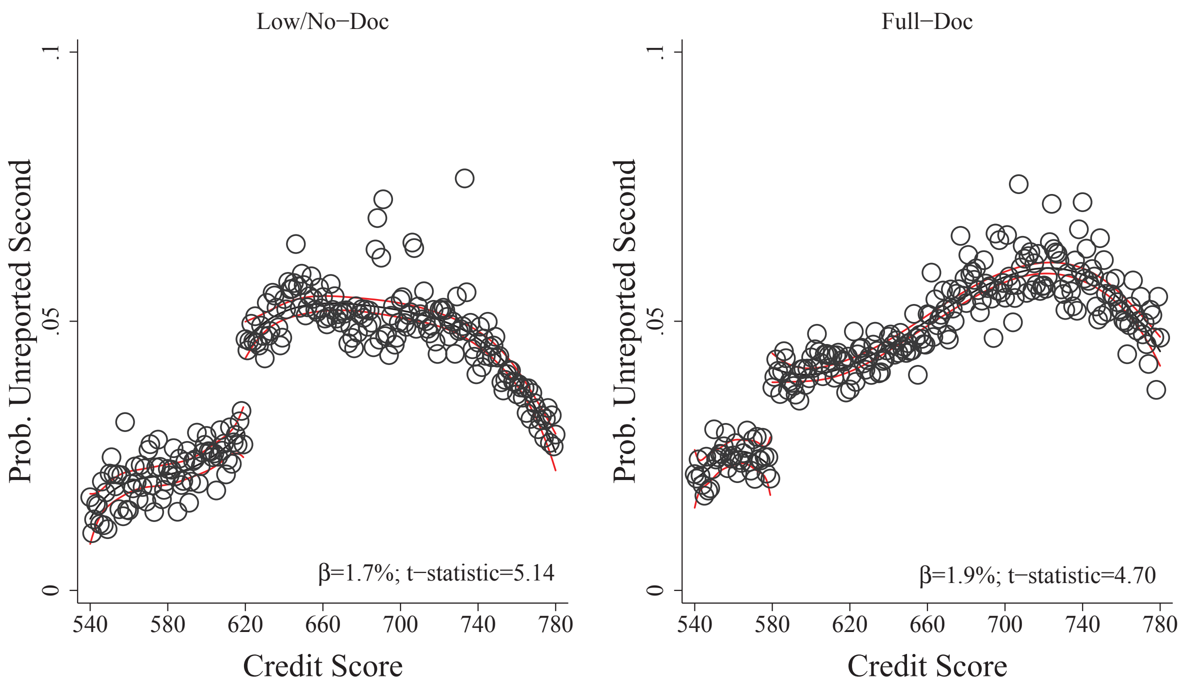


Figure 3. Probability of Second Lien and Occupancy Status Misreporting Around Credit Score Thresholds. This figure shows the probability of having second lien and occupancy status misreporting by credit score. The hollow circles represent the average probability of misreporting for each credit score. The black (dark) line fits a fourth order polynomial approximation at both sides of the credit score threshold (620 for low/no-documentation loans and 580 for full-documentation loans). The red (light) lines delimit the 95% confidence level interval for the approximation. Panel A considers second lien misreporting when the first-second lien pair was either originated by the same or a different lender (all loans). Panels B and C show the result for the decomposition of second lien misreporting. Specifically, Panel B shows the probability of second lien misreporting when the first-second lien pair was originated by the same lender. Panel C shows the probability of second lien misreporting when the first-second lien pair was originated by different lenders. Panel D shows the result for occupancy misreporting. The exact magnitude of the jumps at the discontinuities (β) along with their t-statistics are also shown in the graphs. The corresponding regression results are also shown in Table IA.VII.

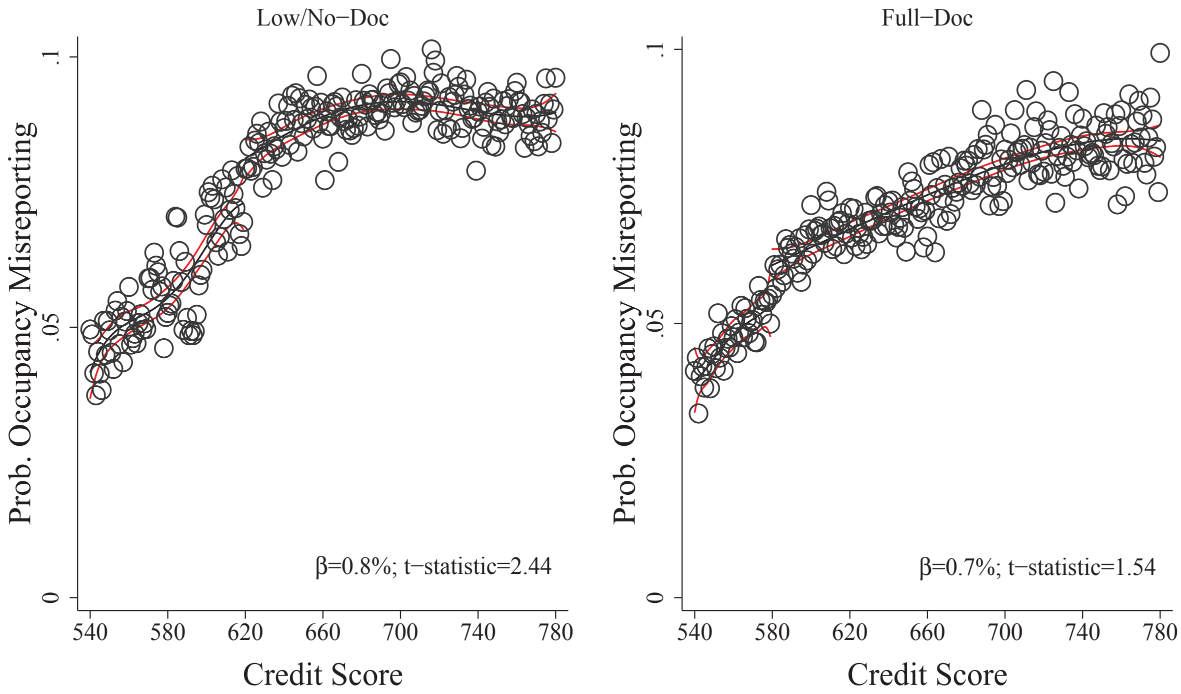
Panel B: Unreported Second (Liens By Same Lender)



Panel C: Unreported Second (Liens By Different Lenders)



Panel D: Occupancy Misreporting



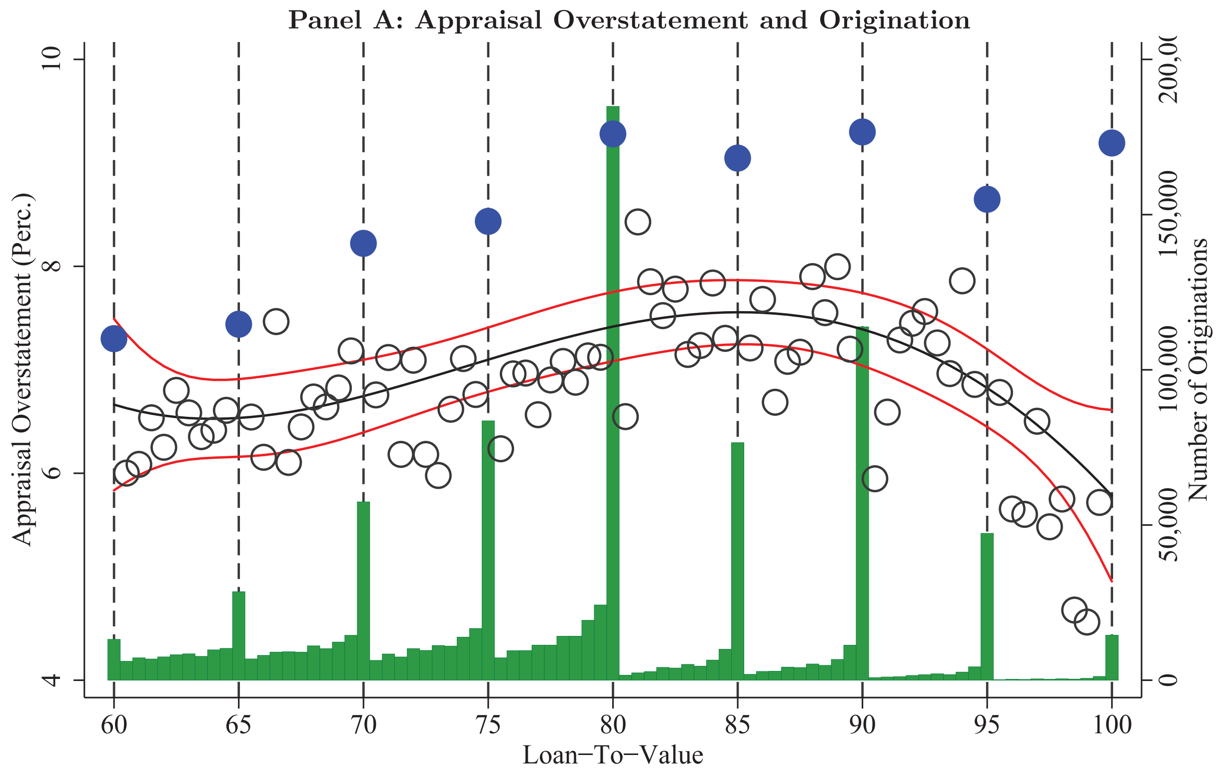
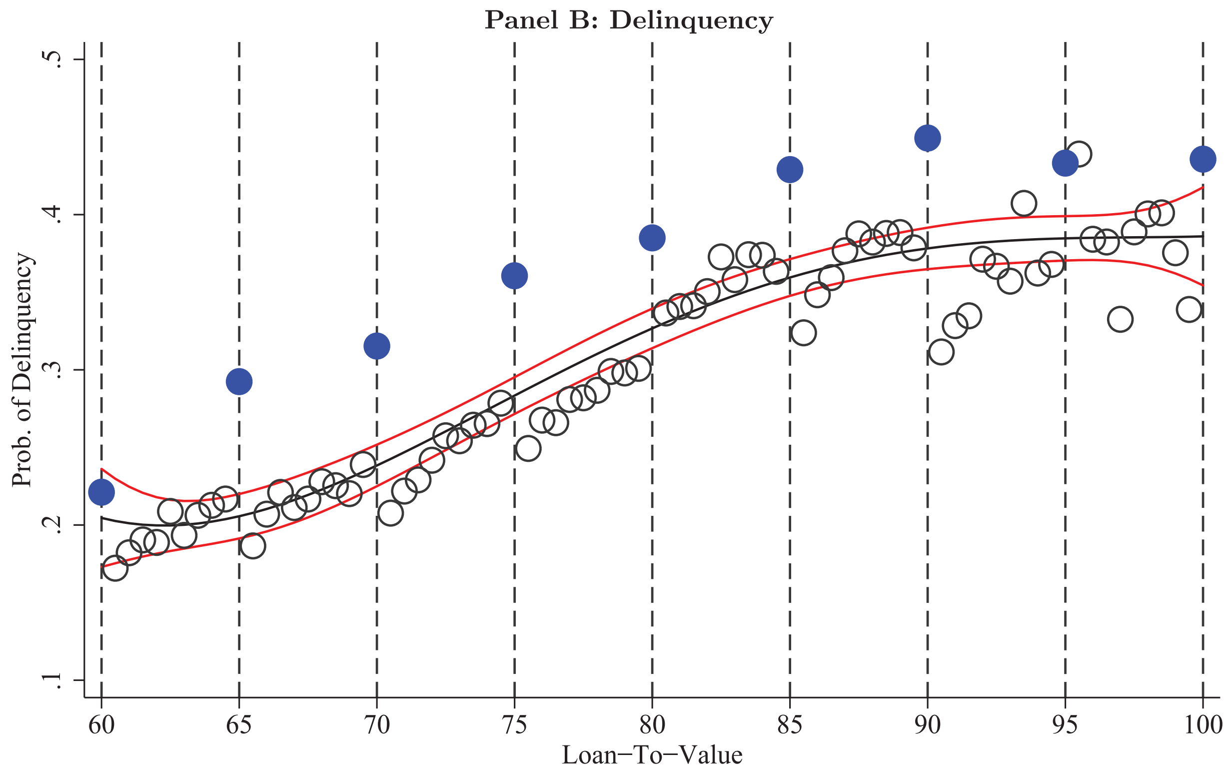
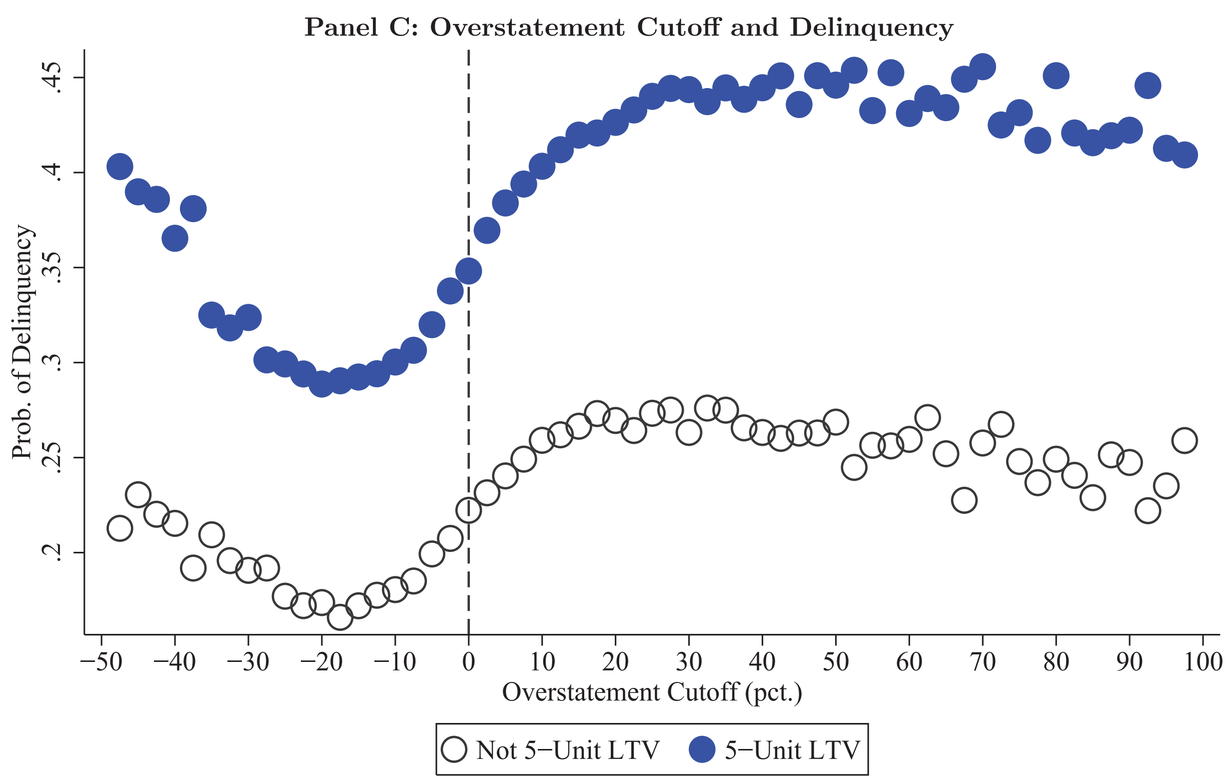


Figure 4. Appraisal Overstatements in Refinances with No Second Liens. This figure shows various features of refinances with no second liens. Panel A shows the mean percentage overstatement of the appraisal compared to the AVM model for each LTV ratio. The black (dark) line fits a fourth order polynomial for appraisal overstatements. The red (light) lines denote the 95% confidence interval. Blue (solid) circles highlight appraisal overstatements at 5-unit LTVs. The green bars show the amount of loan originations by LTV. Panel B shows the probability of delinquency by LTV. As in Panel A, blue (solid) circles highlight delinquency probabilities at 5-unit LTVs. Panel C shows the probability of delinquency for different levels of appraisal overstatements. The blue (solid) circles represent 5-unit LTVs. The black hollow circles represent loans at the remaining LTV levels.


Panel B: Delinquency
Prob. of Delinquency
.1
.2
.3
.4
.5
60
65
70
75
80
85
90
95
100
Loan−To−Value


Panel C: Overstatement Cutoff and Delinquency
Prob. of Delinquency
.2
.25
.3
.35
.4
.45
−50
−40
−30
−20
−10
0
10
20
30
40
50
60
70
80
90
100
Overstatement Cutoff (pct.)
Not 5−Unit LTV
5−Unit LTV

Panel A: Not Controlling By Misreporting

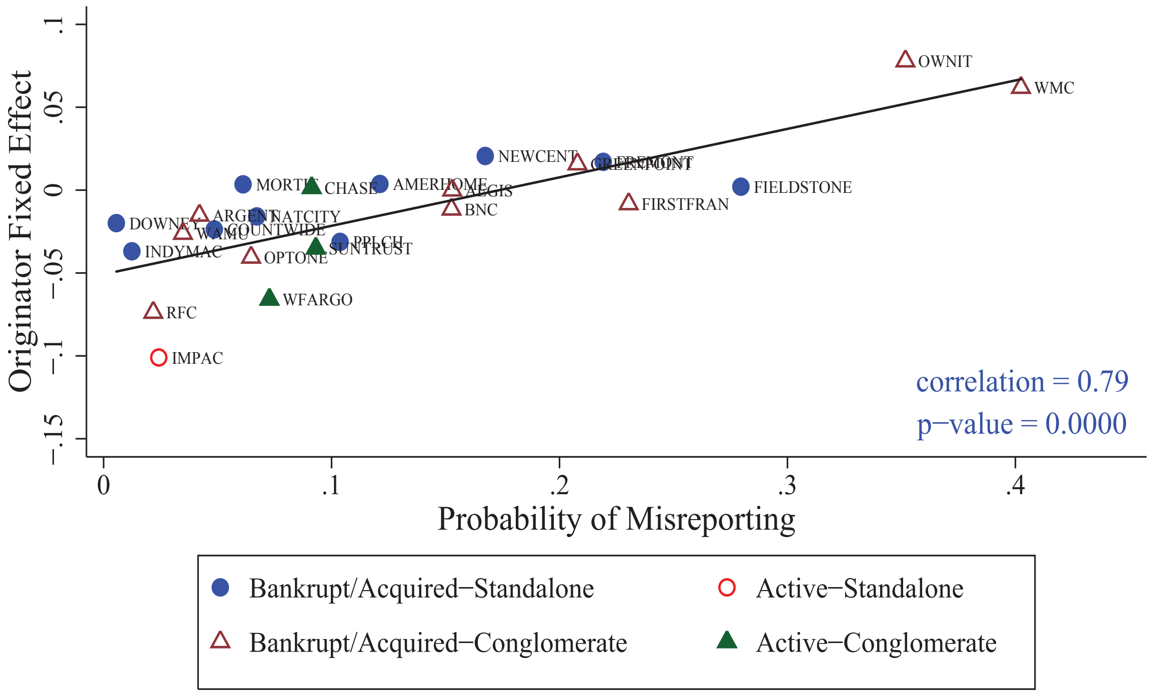
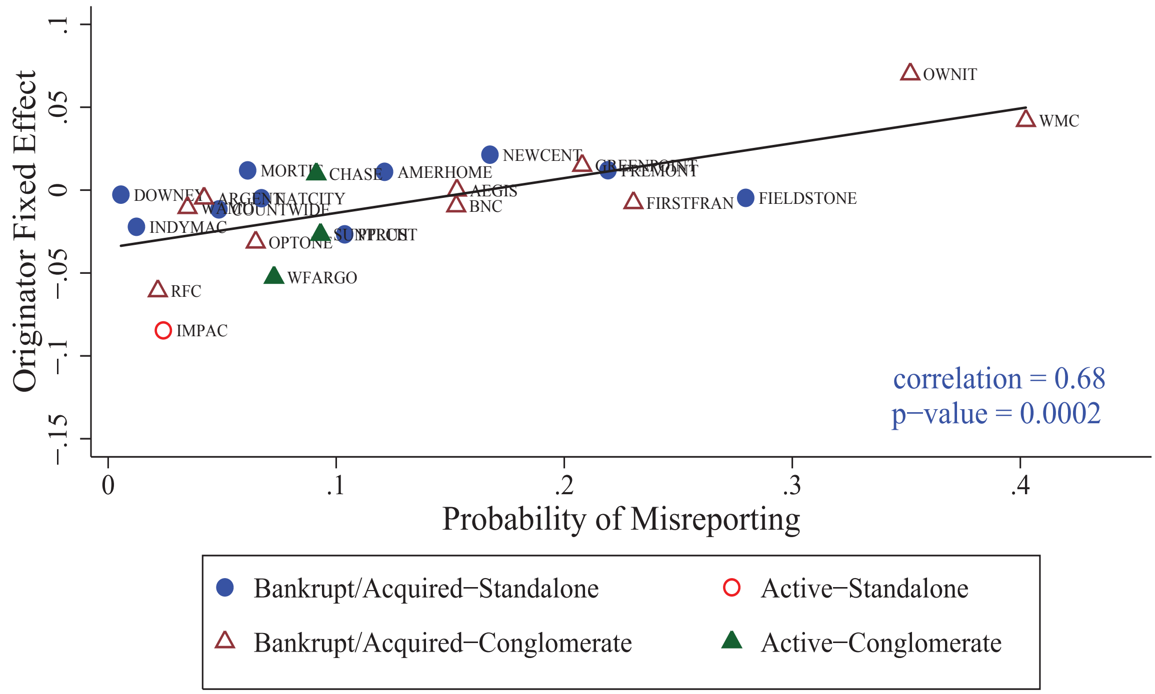


Figure 5. Originator Fixed Effects and Misreporting. We estimate an OLS regression where the dependent variable is the delinquency dummy and the explanatory variables are the set of loan-level controls, CBSA×Quarter fixed effects, and originator fixed effects. Panel A shows the relation between the originator fixed effect estimates and the percentage of loans that exhibit second lien misreporting, by originator. Panel B shows the relation between the originator fixed effect estimate after adding the three potential misreporting indicators (second lien misreporting, occupancy misreporting, and appraisal overstatement) as controls to the previous specification and the percentage of loans that exhibit second lien misreporting, by originator. Blue circles represent originator fixed effects/probability of misreporting pairs of standalone originators that went bankrupt or were acquired during or soon after the recent financial crisis. Hollow red circles correspond to standalone originators that are still active in business. Hollow brown triangles correspond to originators related to a large bank or conglomerate that went bankrupt or were acquired during or soon after the recent financial crisis. Green triangles correspond to originators related to a large bank or conglomerate that are still active in business. The black line fits a linear regression and the correlation is show at the bottom section of each graph. Table IA.XIII in the Internet Appendix presents the names that correspond to each originator's abbreviation.

Panel B: Controlling By Misreporting



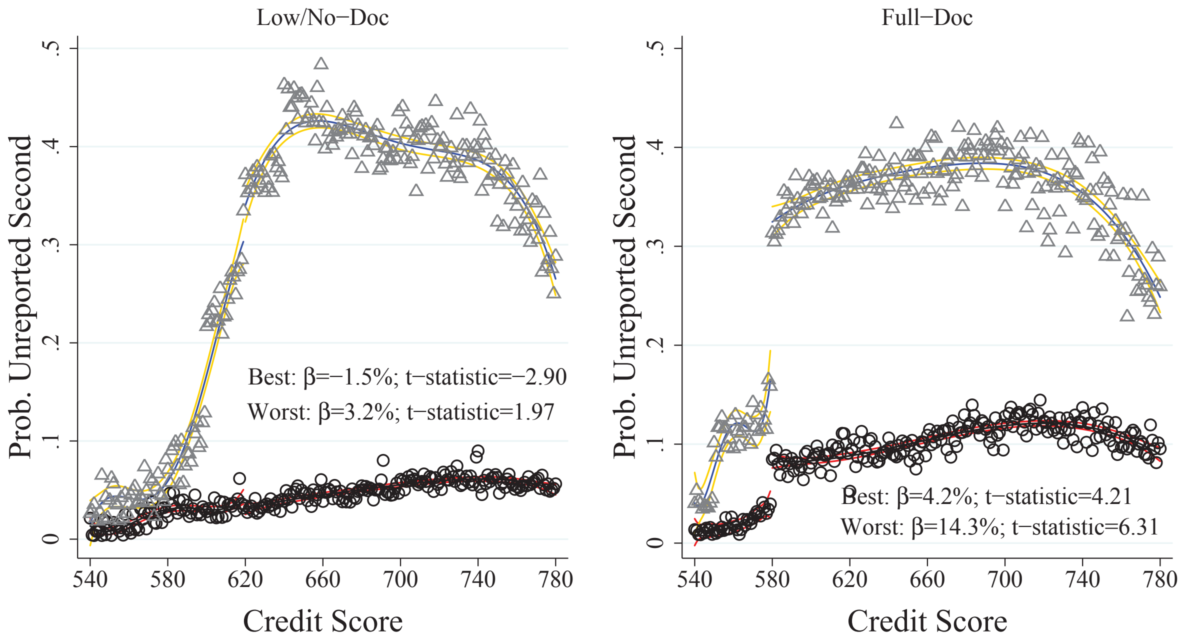


Figure 6. Misreporting Around Credit Score Thresholds, Best and Worst Originators. This figure shows the probability of having second lien misreporting by credit score for the subset of loans originated by the best performing originators and the worst performing originators. The best performers correspond to the tercile of originators with the smallest fixed effect estimates from a regression of the delinquency dummy on loan-level controls, CBSA×Quarter fixed effects, and originator fixed effects. The worst performers correspond to the tercile of originators with the largest fixed effects. We consider the largest 25 originators. The hollow circles represent the probability of second lien misreporting of the best performers for each credit score. The black (dark) line fits a fourth order polynomial approximation at both sides of the credit score threshold (620 for low/no-documentation loans and 580 for full-documentation loans). The red (light) lines delimit the 95% confidence level interval for the approximation. The hollow gray triangles represent the probability of second lien misreporting of the worst performing originators for each credit score. The blue (dark) line fits a fourth order polynomial approximation at both sides of the threshold. The gold (light) lines delimit the 95% confidence level interval for the approximation. The exact magnitude of the jumps at the discontinuities (β) along with their t-statistics are also shown in the graphs. The corresponding regression results are also shown in Table IA.VIII.

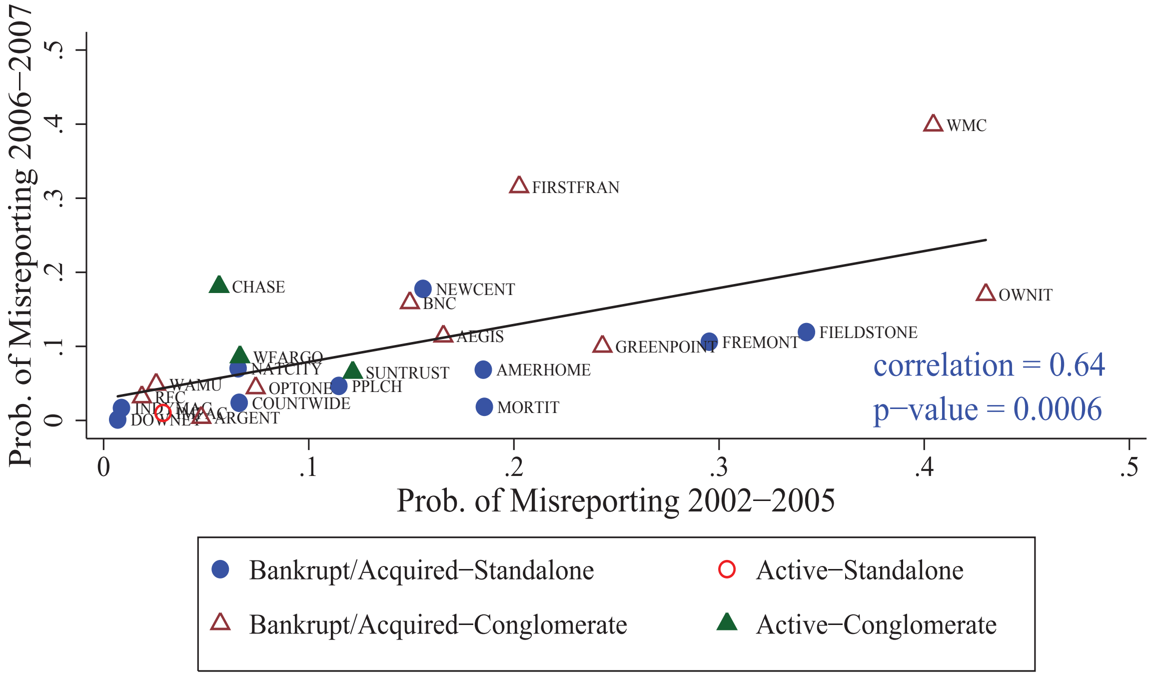


Figure 7. Misreporting Persistence. This figure shows the relation between the amount of second-lien misreporting by originator during the period 2002-2005 and the amount of second-lien misreporting by originator during the period 2006-2007. Blue circles represent standalone originators that went bankrupt or were acquired during or soon after the recent financial crisis. Hollow red circles correspond to standalone originators that are still active in business. Hollow brown triangles correspond to originators related to a large bank or conglomerate that went bankrupt or were acquired during or soon after the recent financial crisis. Green triangles correspond to originators related to a large bank or conglomerate that are still active in business.The black line fits a linear regression and the correlation shown at the bottom section of each graph. Table IA.XIII in the Internet Appendix presents the names that corresponds to each originator's abbreviation.

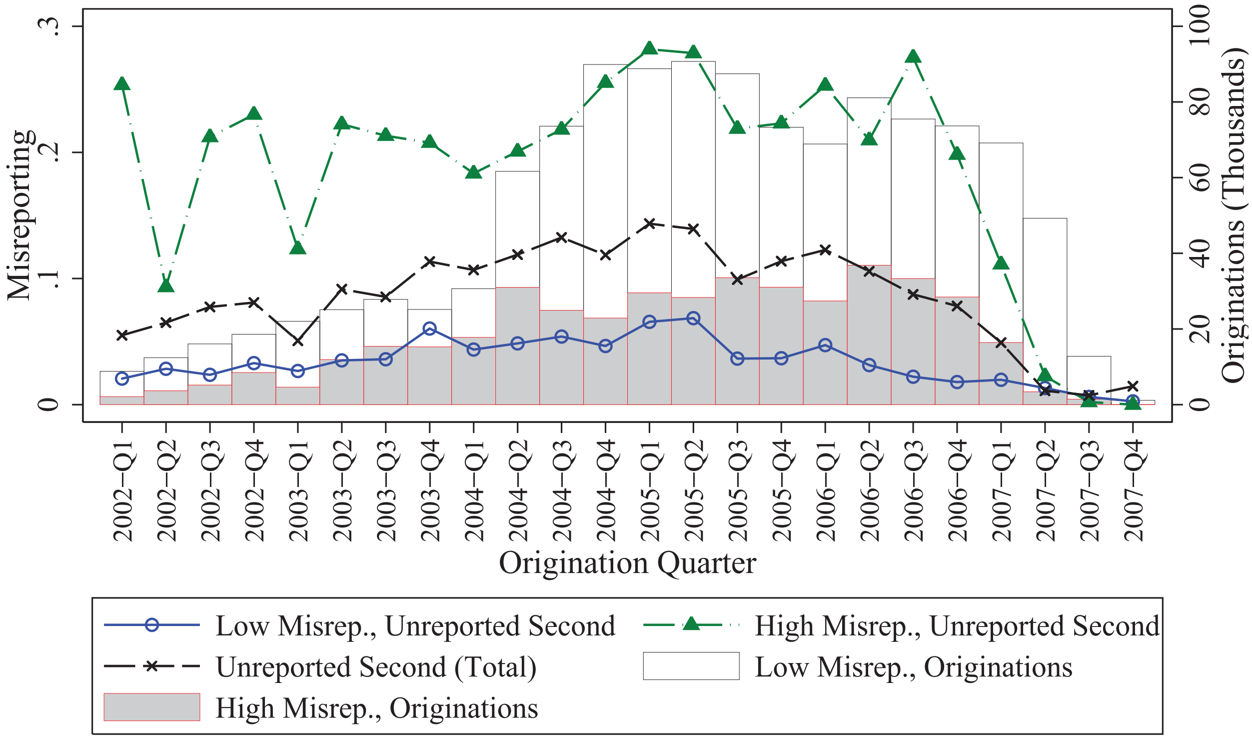


Figure 8. Unreported Seconds by Level of Second Lien Misreporting. We rank originators based on the amount of unreported second liens in 2006. The blue hollow circles show the level of second lien misreporting of the originators in the lowest tercile ("Low Misreporting"). The green triangles show the level of second lien misreporting of the originators in the top tercile ("High Misreporting"). The black dashed line shows the total unreported second liens level in the whole sample (as in Figure 1). The hollow bars and gray bars represent the number of first lien mortgage originations of low-misreporting and high-misreporting originators, respectively.

Panel A: Unreported Second

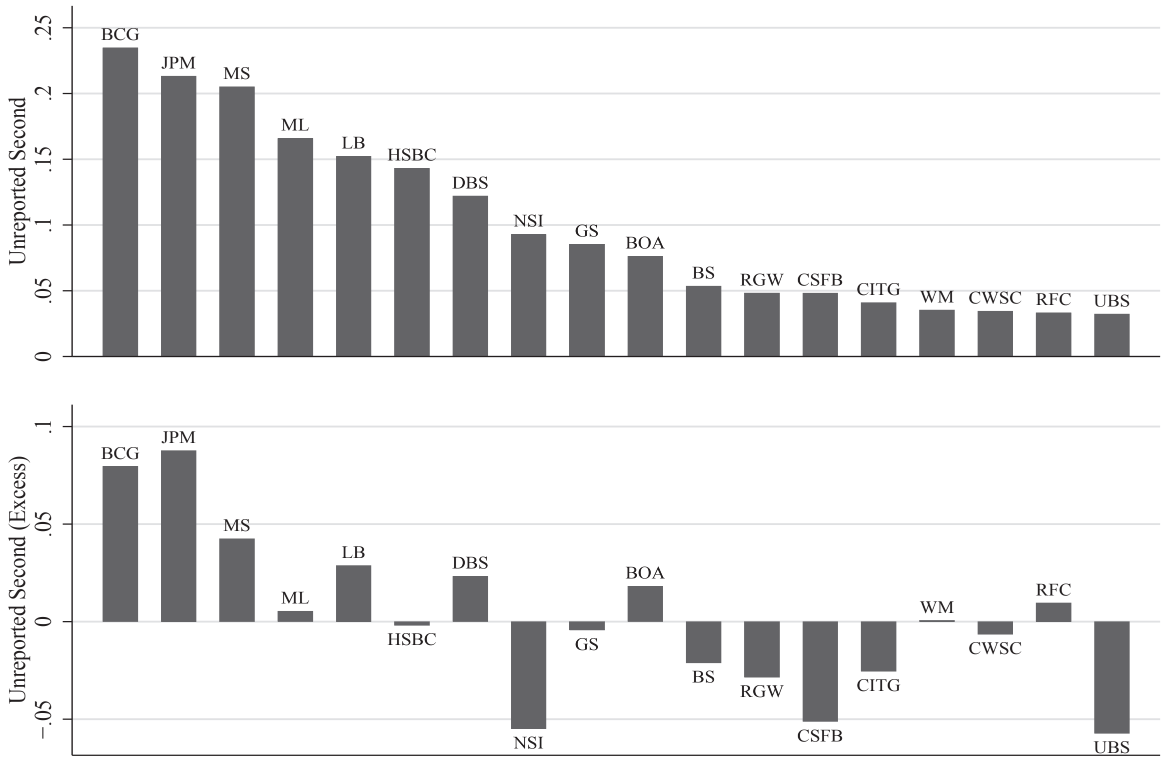
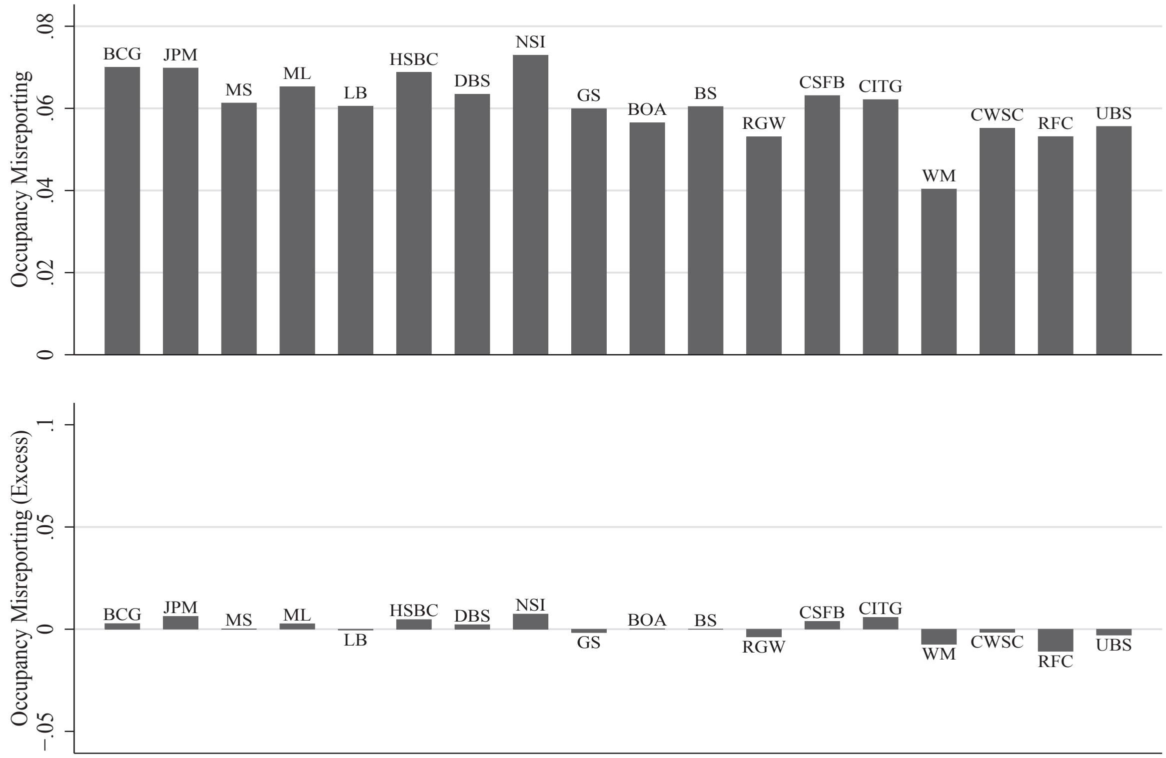
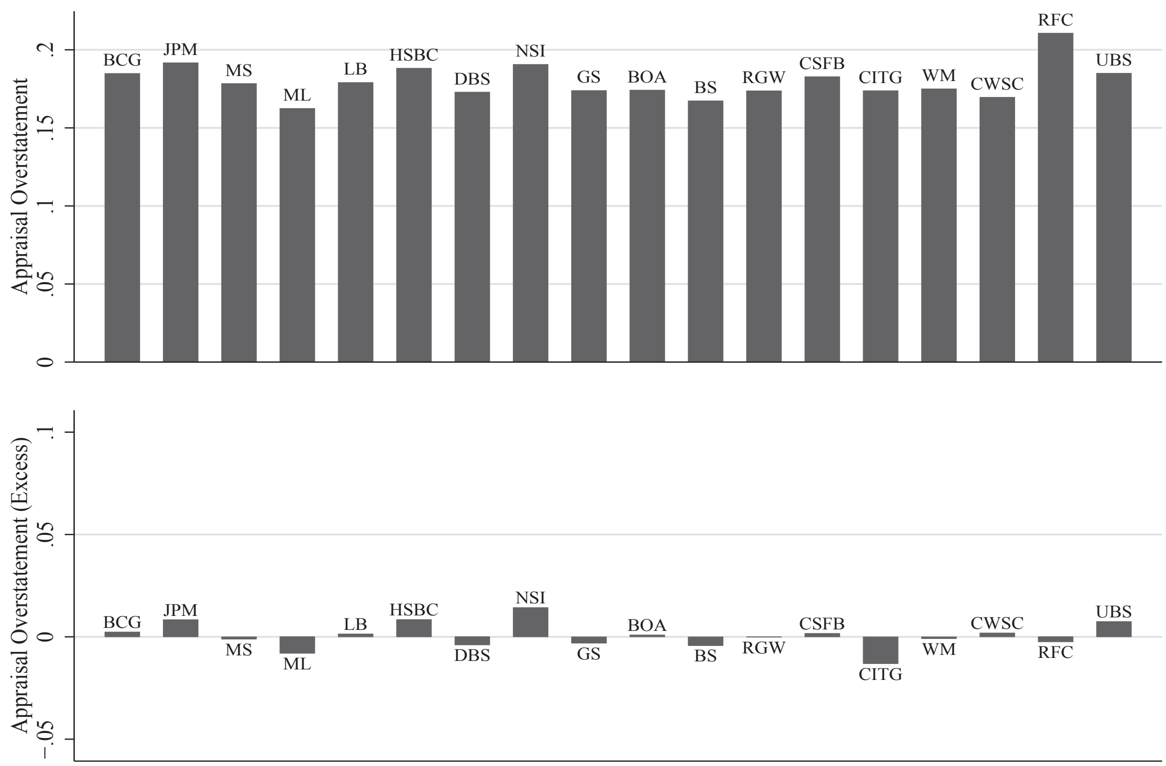


Figure 9. Misreporting By Underwriter. This figure examines differential reporting by underwriters. The top graph in each panel shows misreporting by underwriters. The bottom graph in each panel shows the excess misreporting compared to a benchmark based on the average amount of misreporting of each originator and the importance of each originator in the pools securitized by each underwriter. The comparisons for unreported second liens, occupancy misreporting, and appraisal overstatements are shown in Panels A, B, and C, respectively.

Panel B: Occupancy Misreporting



Panel C: Appraisal Overstatement



Panel D: Misreported

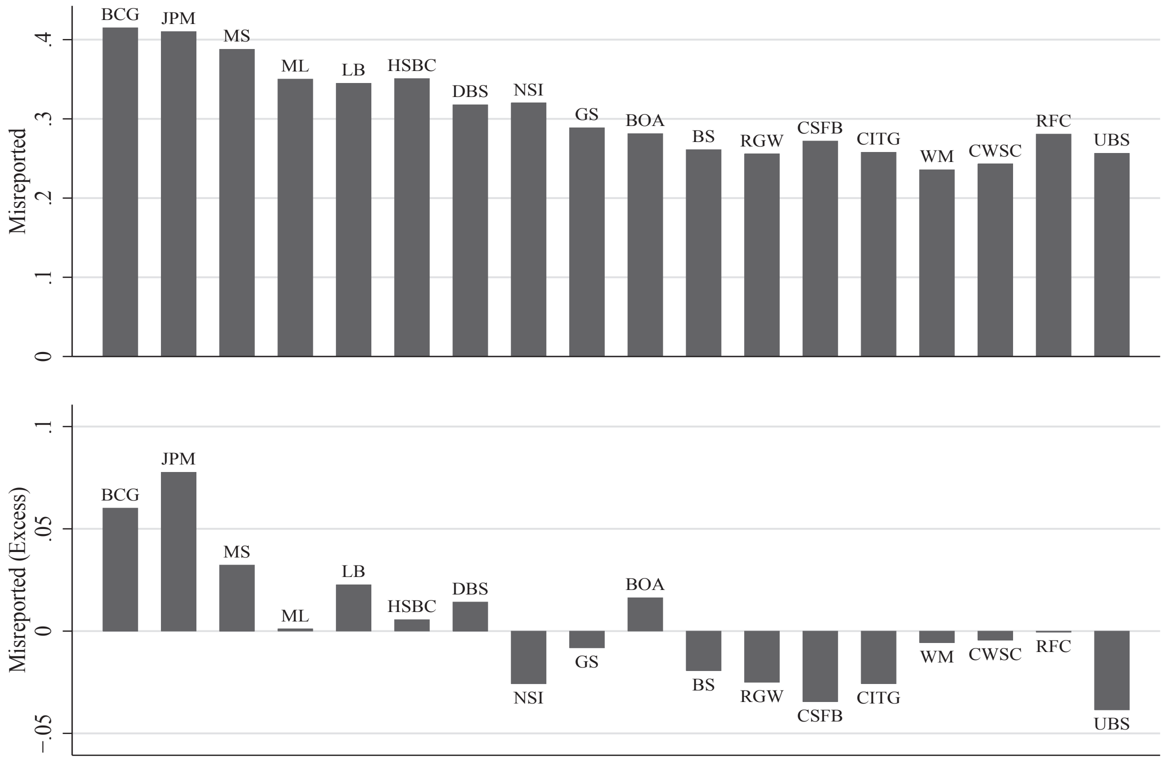


Table I
Sample and Variable Description

This table presents the descriptive statistics and characteristics of the main sample. The sample consists of ABSNet-DataQuick matched securitized first-lien loans that were used for the purchase of a home with an initial loan amount over $30k and a loan-to-value ratio (LTV) lower or equal than 103%. We drop loans associated to the largest 1% of the transactions in each state, loans that are reported as being for homes of more than one unit, and loans that belong to MBS deals in which all mortgages are recorded to have an LTV equal to their combined LTV. Panel A presents descriptive statistics. Panel B presents a description based on different loan characteristics. Panel C describes the performance of the loans.

Panel A: Loan Characteristics			
	Mean	Median	St. Dev
Original Loan Amount	292,721.6	234,000.0	203,532.1
Loan-To-Value	76.7	80.0	13.0
Combined Loan-To-Value	79.8	80.0	14.9
Credit Score	671.5	675.0	73.6
Original Interest Rate	6.6	6.7	2.2

Panel B: Distribution By Characteristics (Pct.)			
Asset Type		*Documentation Type*	
Prime	16.5	Low/No	55.8
Subprime	47.6	Full	44.2
Alt-A, NegArm, Other	36.0	Other/Unreported	-
Loan Purpose		*Reported Occupancy Status*	
Purchase	43.8	Owner Occupied	87.0
Refinance	56.2	Investment/Second Home	13.0
Interest Rate Type		*Prepayment Penalty*	
ARM	67.4	No	38.4
Fixed	32.6	Yes	61.6

Panel C: Loan Performance (Pct.)			
Delinquent	33.1	Direct Default	8.2
Foreclosure	26.1	Early Delinquency	2.2
Number of Loans	3,143,755		

Table II
Misreporting Variables

This table presents the mean values for the misreporting variables (in percent). We construct three potential mortgage misreporting indicators to capture unreported second liens, occupancy status misreporting, and appraisal overstatements. *Unreported Second* is a dummy variable that takes the value of one if the loan in ABSNet associated to that transaction does not disclose the existence of the second lien (i.e., LTV = combined LTV) but both a first and second lien is recorded in county-level recording records as captured in DataQuick. *Occupancy Misreporting* is a dummy variable that takes the value of one if the self-reported occupancy status associated to the loan in ABSNet Loan (using data from MBS prospectus documents) is marked as "owner occupied" but we estimate otherwise from DataQuick's county-level data. *Appraisal Overstatement* is a dummy variable that takes the value of one if the appraised value recorded before origination exceeded the ABSNet's AVM value by more than 20%. *Misreported* is a dummy variable that takes the value of one if one or more of the four misreporting indicators is true.

	Main Sample	Low/No Doc	Full Doc	Purchase	Refinance	Common Sample
Unreported Second	10.2	9.5	11.1	19.6	2.9	8.5
Unreported Second *(From Loans Reported as Having No Second Lien)*	13.4	13.1	13.8	29.3	3.5	11.2
Misreported Occupancy	6.7	7.2	6.0	11.7	2.6	6.0
Misreported Occupancy *(From Loans Reported for Owner Occupancy)*	7.7	8.5	6.7	14.0	2.9	6.8
Appraisal Overstatement	17.8	17.5	18.1	13.2	20.5	17.8
Misreported	30.1	29.6	30.6	35.9	25.6	29.3

Table III
Effect of Misreporting on Delinquency

This table presents the odds ratios of logit regressions where the dependent variable is *Delinquent*, a dummy variable that takes the value of one if the loan was late more than 90 days. Panels A, B, C, and D show the results for unreported second liens, occupancy status misreporting, appraisal overstatements, and aggregated misreporting, respectively. The variables *Unreported Second*, *Occupancy Misreporting*, *Appraisal Overstatement*, and *Misreported* are defined in Table II. The set of controls includes *Reported Second*, a dummy variable that takes the value of one if the loan is reported as having a second lien (i.e., LTV$\neq$ combined LTV) and *Reported Non Owner Occupant*, a dummy variable that takes the value of one if the loan is reported to be for investing in a property or for a second home. Additional controls include controls for low/no-doc loans, loans used for refinancing, borrower's credit score, loan amount, loan-to-value ratio, interest rate at origination, presence of a prepayment penalty, adjustable rate loans, and complex loans. All continuous variables are standardized by subtracting their means and dividing by their standard deviations. All regressions include Core Based Statistical Area (CBSA) times quarter of origination fixed effects (CBSA$\times$Quarter). Standard errors are clustered by CBSA$\times$Quarter of origination. The coefficients for all the additional controls are shown in Table IA.III. z-statistics are presented in parentheses. ***p<0.01, **p<0.05, *p<0.1.

Panel A: Unreported Second

	Main Sample	Main Sample ex CA, FL	Common Sample	Purchases	Refinances
Unreported Second	1.97***	1.97***	2.13***	2.03***	1.81***
	(50.26)	(46.50)	(54.38)	(39.90)	(40.99)
Reported Second	2.14***	2.10***	2.06***	2.63***	1.65***
	(90.00)	(61.29)	(88.55)	(72.93)	(59.84)
Controls	Yes	Yes	Yes	Yes	Yes
CBSA x Quarter FE	Yes	Yes	Yes	Yes	Yes
Observations	3,140,472	1,687,283	2,306,331	1,371,552	1,764,505
Pseudo R^2	0.26	0.21	0.26	0.28	0.25

Panel B: Occupancy Misreporting

	Main Sample	Main Sample ex CA, FL	Common Sample	Purchases	Refinances
Occupancy Misreporting	1.08***	1.14***	1.08***	1.06***	1.05***
	(7.18)	(9.99)	(6.67)	(5.00)	(3.38)
Reported Non Owner Occupied	1.15***	1.30***	1.16***	0.93***	1.46***
	(10.43)	(19.97)	(11.35)	(-3.99)	(30.60)
Controls	Yes	Yes	Yes	Yes	Yes
CBSA$\times$Quarter FE	Yes	Yes	Yes	Yes	Yes
Observations	2,807,954	1,391,821	2,306,331	1,252,643	1,551,027
Pseudo R^2	0.25	0.20	0.25	0.27	0.25

Panel C: Appraisal Overstatement					
	Main Sample	Main Sample ex CA, FL	Common Sample	Purchases	Refinances
Appraisal Overstatement	1.34***	1.36***	1.34***	1.50***	1.23***
	(46.57)	(43.33)	(43.30)	(31.51)	(31.66)
Controls	Yes	Yes	Yes	Yes	Yes
CBSA x Quarter FE	Yes	Yes	Yes	Yes	Yes
Observations	2,576,423	1,334,103	2,306,331	949,997	1,622,278
Pseudo R^2	0.25	0.20	0.25	0.27	0.25

Panel D: Misreported					
	Main Sample	Main Sample ex CA, FL	Common Sample	Purchases	Refinances
Misreported	1.51***	1.53***	1.49***	1.69***	1.30***
	(64.50)	(62.09)	(59.57)	(46.33)	(43.59)
Controls	Yes	Yes	Yes	Yes	Yes
CBSA x Quarter FE	Yes	Yes	Yes	Yes	Yes
Observations	3,141,156	1,687,666	2,306,331	1,371,613	1,765,126
Pseudo R^2	0.25	0.21	0.26	0.28	0.25

Table IV
Effect of Misreporting on the Interest Rate at Origination

This table presents OLS regressions where the dependent variables is the interest rate at origination. Panels A, B, and C show the results for unreported second liens, occupancy status misreporting, and appraisal overstatements, respectively. The variables *Unreported Second*, *Occupancy Misreporting*, and *Appraisal Overstatement* are defined in Table II. The set of controls includes *Reported Second*, a dummy variable that takes the value of one if the loan is reported as having a second lien (i.e., LTV$\neq$ combined LTV) and *Reported Non Owner Occupant*, a dummy variable that takes the value of one if the loan is reported to be for investing in a property or for a second home. Additional controls include controls for low/no-doc loans, loans used for refinancing, borrower's credit score, loan amount, loan-to-value ratio, interest rate at origination, presence of a prepayment penalty, adjustable rate loans, and complex loans. LTV is separated into two components: values of 80 or lower (*LTV Low*) and values over 80 (*LTV High*). All continuous variables are standardized by subtracting their means and dividing by their standard deviations. All regressions include Core Based Statistical Area (CBSA) times quarter of origination fixed effects (CBSA$\times$Quarter). Standard errors are clustered by CBSA$\times$Quarter of origination. The coefficients for all the additional controls are shown in Table IA.VI. *t*-statistics are presented in parentheses. ***p<0.01, **p<0.05, *p<0.1.

Panel A: Unreported Second

	Main Sample	Main Sample ex CA, FL	Common Sample
Unreported Second	0.14***	0.10***	0.14***
	(19.35)	(11.48)	(15.33)
Reported Second	0.10***	0.06***	0.09***
	(11.31)	(7.72)	(9.81)
Controls	Yes	Yes	Yes
CBSA x Quarter FE	Yes	Yes	Yes
Observations	3,140,472	1,687,283	2,306,331
Adj. R^2	0.62	0.60	0.62

Panel B: Occupancy Misreporting

	Main Sample	Main Sample ex CA, FL	Common Sample
Occupancy Misreporting	0.05***	0.07***	0.04***
	(9.73)	(11.19)	(8.51)
Reported Non Owner Occupied	0.32***	0.42***	0.32***
	(45.54)	(54.41)	(43.51)
Controls	Yes	Yes	Yes
CBSA$\times$Quarter FE	Yes	Yes	Yes
Observations	2,807,954	1,391,821	2,306,331
Adj. R^2	0.63	0.60	0.62

Panel C: Appraisal Overstatement

	Main Sample	Main Sample ex CA, FL	Common Sample
Appraisal Overstatement	0.07***	0.09***	0.07***
	(25.37)	(23.82)	(23.52)
Controls	Yes	Yes	Yes
CBSA x Quarter FE	Yes	Yes	Yes
Observations	2,576,423	1,334,103	2,306,331
Adj. R^2	0.62	0.60	0.62

Table V
Determinants of Misreporting

This table compares the adjusted R-Squares obtained from OLS regressions of the misreporting indicators on loan-level controls and different combinations of originator and underwriter fixed effects. Underwriter information comes from Bloomberg. Only loans belonging to MBS deals where one main underwriter is listed are considered. The variables *Unreported Second*, *Occupancy Misreporting*, and *Appraisal Overstatement* are defined in Table II. The complete regressions are shown in Table IA.IX.

	Unreported Second		Occupancy Misreporting		Appraisal Overstatement	
	Adj. R^2	Relative Increase (Perc.)	Adj. R^2	Relative Increase (Perc.)	Adj. R^2	Relative Increase (Perc.)
Baseline	0.100	-	0.087	-	0.085	-
w/ Originator FE	0.152	52.2	0.089	1.4	0.086	1.7
w/ Underwriter FE	0.129	29.1	0.087	0.1	0.085	0.4
w/ Originator and Underwriter FE	0.168	68.2	0.089	1.5	0.086	1.8
Sample Size	1,726,075					

Table VI

Appraisal Overstatements in Refinances

This table shows the odds ratios of logit regressions where the dependent variable are *Appraisal Overstatement* or *Delinquent*. Refinances are split into cashout refinances and term refinances. *Appraisal Overstatement* $\times$ *LTV* 5 captures appraisal overstatements for loans that have 5-unit LTVs. The regression controls for reported second liens, reported non owner occupied, low/no-doc loans, loans used for refinancing, borrower's credit score, loan amount, loan-to-value ratio, interest rate at origination, presence of a prepayment penalty, adjustable rate loans, and complex loans. All regressions include Core Based Statistical Area (CBSA) times quarter of origination fixed effects (CBSA$\times$Quarter). Standard errors are clustered by CBSA$\times$Quarter of origination. The complete regressions are shown in Table IA.X. z-statistics are presented in parentheses. ***p<0.01, **p<0.05, *p<0.1.

Panel A: Appraisal Overstatement and Refinances		
	Appraisal Overstatement	
	Main Sample	Main Sample
Refinance	1.74***	
	(39.56)	
Cashout Refinance		1.81***
		(42.09)
Term/Rate Refinance		1.57***
		(28.80)
CBSA x Quarter FE	Yes	Yes
Observations	2,575,484	2,560,060
Pseudo R^2	0.09	0.09

Panel B: Delinquencies and Appraisal Overstaements		
	Delinquent	
	Main Sample	Main Sample
Appraisal Overstatement x LTV5		1.25***
		(22.85)
Appraisal Overstatement	1.34***	1.16***
	(46.57)	(15.68)
CBSA$\times$Quarter FE	Yes	Yes
Observations	2,576,423	2,576,423
Pseudo R^2	0.25	0.25

Table VII
Effect of Misreporting on MBS Losses

This table summarizes the results of the deal-level regressions. The dependent variable is *ln(1 + Losses)*. The explanatory variables consist on the aggregated misreporting by deal and the aggregated deal-level controls. All regressions have Underwriter×Vintage fixed effects and standard errors are clustered by Underwriter×Vintage. Panel A shows the effect of misreporting on MBS losses. Panel B shows the effect of the other controls. The complete regressions are shown in Table IA.XII and the descriptive statistics of the deal variables are shown in Table IA.XI. *t*-statistics are presented in parentheses. ***p<0.01, **p<0.05, *p<0.1.

Panel A: Effect of Misreporting

	Unreported Second	Occupancy Misreporting	Appraisal Overstatement	Misreported
Standard Deviation (Perc.)	12.7	2.3	4.4	10.4
OLS Estimate	0.08***	0.39***	0.25***	0.15***
Increase in Loss (Perc.)	1.18	1.04	1.28	1.82
Mean Loss	15.9			
Increase in Loss Relative to Mean (Perc.)	7.4	6.6	8.1	11.5

Panel B: Effect of Other Variables

	Standard Deviation	OLS Estimate	Increase in Loss (Perc.)
Reported Second (Perc.)	19.8	0.11***	2.55
Reported Non-Owner Occupied (Perc.)	13.9	0.02	0.32
Low/No-Doc (Perc.)	25.2	0.04***	1.18
Refinance (Perc.)	16.2	0.04***	0.75
Mean Creditscore	46.8	-0.00048***	-2.57
Mean LTV	5.4	0.00025	0.16
Prepayment Penalty (Perc.)	30.7	0.07***	2.52
Complex (Perc.)	31.8	0.05***	1.86
ARM (Perc.)	41.4	0.03***	1.45
Mean Interest Rate	1.9	0.01***	2.17
Second Lien (Perc.)	7.2	0.08*	0.67
Mean Loan Amount	158,025	-0.00000009***	-1.64
Number of Loans	3,212	-0.00000102***	-0.38